1993-12

Applying theory of constraints to an aircraft remanufacturing line

Beck, James Michael.

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/39658

(2)

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

APPLYING THEORY OF CONSTRAINTS TO AN AIRCRAFT REMANUFACTURING LINE

by

James Michael Beck

December 1993

Principal Advisor: Shu S. Liao

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
December 1993

3. REPORT TYPE AND DATES COVERED
Master's Thesis

4. TITLE AND SUBTITLE
APPLYING THEORY OF CONSTRAINTS TO AN AIRCRAFT REMANUFACTURING LINE

5. FUNDING NUMBERS

6. AUTHOR(S)
Beck, James M.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Theory of Constraints (ToC) is a production philosophy seeking to globally optimize the manufacturer's constraints to meeting his business objective. The constraints to the Department of Defense (DOD) Naval Aviation Depot (NADEP) are: Unknown production resource requirements until aircraft disassembly and unscheduled production workloads due to unexpected high unit operational failure or unit modification. The thesis objective is to analyze the NADEP's remanufacturing process, determine areas that will benefit from the implementation of ToC, and then document managerial tools the depot level manager could use to globally optimize the remanufacturing process. The study focuses on the A-6 remanufacturing line at the NADEP, Alameda, CA. The thesis concludes ToC is a viable methodology, and with the use of Queuing Charts and Critical Resource Scheduling, the NADEP can meet its business objective of providing a cost effective remanufacturing capability to the DOD.

14. SUBJECT TERMS
Theory of Constraints, Aircraft Remanufacturing, NADEP, Critical Resource Scheduling, Queuing Theory

15. NUMBER OF PAGES
90

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std Z39-18

Approved for public release; distribution is unlimited.

Applying Theory of Constraints to an Aircraft Remanufacturing Line

by

James M. Beck
Lieutenant Commander, United States Navy
B.A., University of California-Los Angeles, 1979

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
December 1993

Author: [redacted]

James M. Beck

Approved by: [redacted]

Shu S. Liao, Principal Advisor

[redacted]

Martinus Sarigul-Klijn, Associate Advisor

[redacted]

David R. Whipple, Chairman
Department of Administrative Sciences

ABSTRACT

Theory of Constraints (ToC) is a production philosophy seeking to globally optimize the manufacturer's constraints to meeting his business objective. The constraints to the Department of Defense (DOD) Naval Aviation Depot (NADEP) are: Unknown production resource requirements until aircraft disassembly and unscheduled production workloads due to unexpected high unit operational failure or unit modification. The thesis objective is to analyze the NADEP's remanufacturing process, determine areas that will benefit from the implementation of ToC, and then document managerial tools the depot level manager could use to globally optimize the remanufacturing process. The study focuses on the A-6 remanufacturing line at the NADEP, Alameda, CA. The thesis concludes ToC is a viable methodology, and with the use of Queuing Charts and Critical Resource Scheduling, the NADEP can meet its business objective of providing a cost effective remanufacturing capability to the DOD.

TABLE OF CONTENTS

ACKNOWLEDGEMENT

Thanks to Commander Marti Sarigul-Klijn for his scientific innovative insight into the remanufacturing process. Without his vision, production goals would continue to be short ranged and only locally optimizing. His strength to make change happen is one of his best attributes.

I. INTRODUCTION

A. BACKGROUND

The Department of Defense (DOD), during World War II, needed massive civilian corporate support for the armed forces. After the war, the DOD decided it needed to have its own industrial base to support the armed forces. This led to the national policy, "to provide a ready and controlled industrial base for mobilization which is also cost effective in peace times." [Ref.1:p.8]

The primary function of the DOD industrial base is to provide a remanufacturing capability of military equipment damaged from war or needing repair because of age. Remanufacturing is different from most forms of manufacturing because of:

- Unknown production resource requirements until disassembly of the unit needing remanufacturing.
- Unscheduled production workloads due either to unexpected high unit operational failure or to unexpected unit modification requirements implemented after the unit's induction to the remanufacturing process.

These differences are also constraints affecting the way these DOD activities are able to successfully conduct their mission of providing needed, high quality, cost effective, remanufacturing support.

Unknown production resource requirements and unscheduled production workloads have the combined effect of causing large fluctuations in the work demand of the activity. If the production rate does not keep up with fluctuations in the customers work demand rate, then the work in process (WIP) accumulates. The Naval Aviation Depot (NADEP) is responsible for the remanufacturing of military aircraft and its component parts. There are six such NADEPs in DOD. The NADEP at Alameda, Ca., in particular, has also been unable to keep its production rate matched with the aircraft work demand rate of its customers. On the A-6 remanufacturing line, this has resulted in a large accumulation of aircraft in process (AIP).

The A-6 line suffered an $18 million loss (one million per aircraft) last fiscal year because high levels of AIP resulted in high operating expenses. Work turnaround time (TAT) steadily increased from slightly over one half year to several calendar years because limited production resources were stretched over higher AIP. The increased TAT also caused the

NADEP to have poor compliance with the completion schedule forecasted for its customers. High AIP also generated quality control problems at the NADEP.

These four obstacles to being a successful aircraft remanufacturing activity, high AIP, high operating expenses, high TAT, and poor quality control, are not unique to NADEP, Alameda. They are also problems in part or in whole of all six NADEPs.

To address the problems facing the NADEP remanufacturing operation, this study will examine the feasibility of applying modern manufacturing tools such as Just-In-Time, queuing, and Theory of Constraints (ToC). Since ToC closely resembles the JIT methodology and seems to provide a more generalizable approach to various manufacturing environments, this study will focus on the Theory of Constraints rather than JIT.

ToC is a production philosophy, evolved by Dr. Eliyahu M. Goldratt, seeking to globally optimize all the manufacturer's constraints to meeting the business objective. ToC seeks to reduce accumulated WIP and decrease TAT, increase schedule compliance and improve meeting financial goals, and also improve quality and managerial control.

B. OBJECTIVE

The thrust of the research was to conduct an investigation on the remanufacturing process in NADEP and analyze the feasibility of applying ToC as an alternative remanufacturing method in the hope of alleviating the problems discussed earlier.

Data from the A-6 line at NADEP, Alameda is used to illustrate the application of ToC in the remanufacturing process.

C. RESEARCH QUESTIONS

The primary research question was "Is ToC applicable to the NADEP aircraft remanufacturing environment?" The subsidiary research questions were:

- What are the production principles of ToC and how are they practically defined?
- What tool can the aviation depot manager use to better monitor inductions and completions?
- How can aviation depot managers better schedule critical resources?

D. SCOPE, LIMITATIONS AND ASSUMPTIONS

The study focused on the A-6 remanufacturing line at the NADEP, Alameda, CA. An analysis was conducted of the

remanufacturing line to determine areas that could benefit from the implementation of ToC philosophy. Although there are other aircraft remanufacturing lines at the NADEP, the study was limited to the A-6 line because of the similarity between various aircraft remanufacturing lines.

E. ORGANIZATION OF THE THESIS

The following chapter will describe the A-6 remanufacturing line at the NADEP. Chapter III explains specific topics of Queuing Theory that will help in Chapter IV to define the relevant production principals of ToC. Chapter IV discusses ToC as a logical approach to global optimization in the remanufacturing environment. Chapter V provides a blueprint for implementation of ToC at the A-6 remanufacturing line, provides insight into scheduling critical resources, and presents the use of queuing charts as a management tool. Chapter VI provides a summary of information developed in the thesis, listing conclusions, recommendations and areas for follow-on research.

Appendix A is a Lotus 123 spreadsheet program to allow the plant manager to develop the queuing chart tool. Appendix B

is a resource planning problem. It will illustrate Critical Resource Scheduling (CRS) method.

II. A-6 AIRCRAFT REMANUFACTURING LINE

The A-6 remanufacturing process of NADEP-Alameda separates into four different lines:

- **A-6 Standard Depot Level Maintenance (SDLM).** The material condition of the aircraft at induction dictates the level of rework necessary.

- **A-6 SDLM Rewing.** These aircraft receive both the SDLM rework and the new composite wing structure.

- **A-6 SDLM 601.** Airframe modification 601 is an interim air frame change to increase the life of the wings. There were only six such aircraft inducted at Alameda before it was proven that the change would not work.

- **EA-6B** *SDLM. Same as the A-6 SDLM except on the four seat* EA-6B aircraft.

This chapter looks at the remanufacturing line in both the macro and micro level. It discusses dependent events, statistical fluctuations and current management control systems as they relate to the remanufacturing process.

A. DEPENDENT EVENTS

Dependent events are a series of operations, each one performed before the subsequent "downstream" operation.

1. Macro Level

The assembly line is the most basic depot structure. It became famous when Henry Ford first built cars in large volume. It is one continuous line of dependent events and hence called an "I" structure because of the physical similarity to the letter I. The dependent events in this structure are easy to recognize because each consecutive process connects to form one long, continuous conveyor belt.

The A-6 remanufacturing line, at the macro level, is an "I" structure as depicted in Figure 1. The first event is acceptance and paint strip. The "40" shop examines and evaluates the material condition of the installed engines, flight controls, aircraft structure, electrical and hydraulic installations, and related systems and components. The "40" shop also examines the aircraft for technical directives compliance/incorporation. It then documents the depth of work to be accomplished. The "40" shop also determines material and configuration requirements, and provides detailed rework instructions. Shops 20 - 26 are responsible for disassembly, repair and modification of aircraft structures. Shops 31 - 34 are responsible for reassembly of mechanical and electrical

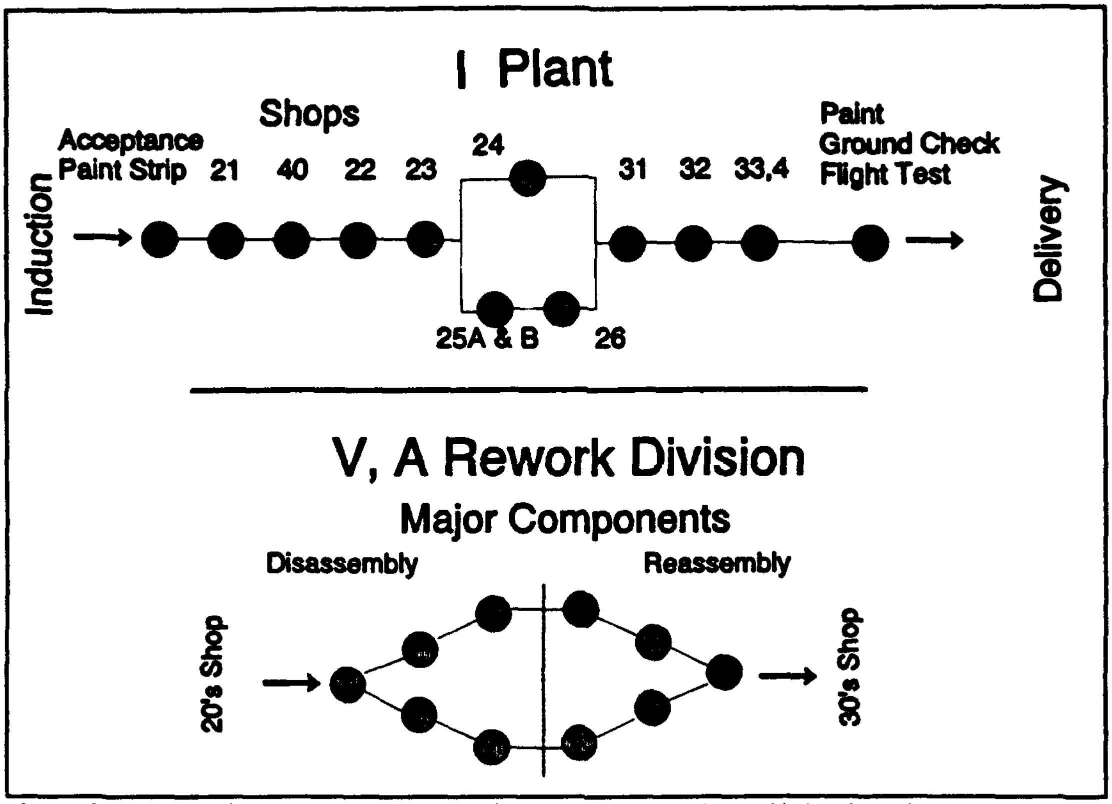


Figure 1: The serial, dependent "I" logical structure and parallel, dependent V, A logical structure of the A-6 remanufacturing line.

components, flight control rigging, operational/functional checks, systems tests, and the final closeout of the remanufactured aircraft. The operations support division is responsible for painting and preparing the aircraft for flight. The final event is the test flight and delivery of the aircraft to the customer.

Using the "I" logical structure allows for simpler work breakdown by the Aircraft Planning Branch, the "10" shop. The Aircraft Planning Branch determines workload commitments for induction and scheduling of aircraft and sets priorities for

resource utilization. It also develops and maintains the Master Rework Plan. The Master Rework Plan specifies the extent of aircraft disassembly required, sequence of operations, and in-process times. The "10" shop makes workload commitments for induction and scheduling, sets priorities and obtains materials to support the rework of aircraft and accessories.

The "I" logical structure at the macro level is a series of dependent events from acceptance, paint strip, examination and evaluation, disassembly, structural repair, modification, reassembly, paint, ground check, flight check, and delivery. However, at the micro level, there is more than just the "I" structure. After examination, the removed airframe components go to rework division shops outside the macro view of the A-6 line.

2. Micro Level

The micro view of the remanufacturing line shows rework divisions branching off the "I" structure. The rework divisions include Avionics, Landing Gear, Components, Ground Service Equipment, Manufacturing, and Power Plants. After removing major components of the aircraft by the 20's shops

and inspecting by the 40 shop, the rework divisions further break down the components into sub-components. The sub-components are reworked, reassembled into components and given to the 30's shop on the "I" structure, aircraft line.

Figure 1 presents each rework division as a "V" logical structure attached to an "A" logical structure. The "V" logical structure contains the diverging, disassembly and repair work center production points in the remanufacturing process. The "A" logical structure contains the converging, reassembly work center production points in the remanufacturing process. The reassembly shops, shops 31 and 34, on the "I" structure are dependent on the output of the rework divisions.

Components sent to the rework divisions are of two types, interchangeable and matched. Matched components require returning to a particular major component or aircraft. Interchangeable components exchange with other like components. If matched components from the rework divisions are not ready in time, a backlog of work at the reassembly "I" structure shop begins to grow. This backlog may cause the whole process to halt.

B. STATISTICAL FLUCTUATIONS

The process time of A-6 aircraft through the remanufacturing line has an exponential distribution because of the statistical fluctuations in the production flow. As the production flow increases, the fluctuations in the flow increase exponentially.

These fluctuations are caused by:

- **Induction Rate.** The time interval between successive aircraft arrivals is independent of other arrivals and is not constant in time.
- **Variance in Work Content of Each Aircraft.** Each aircraft has a tailored rework process assigned to it. The level of rework is not accurately determined before induction because every aircraft is different.
- **Late Finds.** Field teams conduct a pre-inspection of the aircraft before delivery to the NADEP. The field teams get a rough idea of the level of rework the aircraft will need. However, upon doing a complete breakdown of the aircraft inducted, additional rework requirements may emerge.
- **Engineering Changes on Aircraft in Process.** Either scheduled or unscheduled engineering changes occur during the rework process. Unscheduled changes can cause a serious constraint at the responsible shop even if it has excess service capacity. Scheduled engineering changes can cause fluctuation in process flow.
- **Material Shortages.** Not having the proper materials available when required stops the process flow.
- **Key Personnel not Available.** This is the same situation as having a material shortage.

- **Large Waiting Times Due to High Work in Process (WIP).** Accumulating WIP slows down the work flow because labor resources become spread thin.

C. CURRENT MANAGEMENT OF AN AIRCRAFT REMANUFACTURING LINE

The remanufacturing process is a labor intensive production environment. Manpower is a constraint. Also, the cost of manpower is very high and one of the largest contributors to operating expense. From an accounting standpoint, if labor in a shop is not producing, the efficiency of that shop will be low. The inclination of management is make the shop produce more to keep the efficiencies high. This production local optimization causes a backlog of work in two locations: At the shop with the original low efficiences and second, at the downstream shop that might already have high efficiencies but now is being overloaded with work. The next two chapters will show why local optimization causes long turnaround time, low schedule compliance and poor quality control.

III. QUEUING THEORY

A. INTRODUCTION

The purpose of this chapter is to explain specific topics of Queuing Theory that will define the relevant production philosophies of Theory of Constraints (ToC) in Chapter IV.

No one directly benefits from queuing. The time a customer spends waiting in a line is time wasted, it does not benefit the customer or the organization. The information presented in this chapter is from Queuing Methods, by Randolph Hall and tailored to the remanufacturing environment.

B. DEFINITIONS

1. Elements of Queuing Systems

There are three basic elements in a typical queuing system:

- Customer. The person or object that waits for service. At the Naval Aviation Depot (NADEP), an aircraft or component part is a customer.
- Server. The object providing the service. The Depot itself, as well as the artisan who does the rework, are both servers. A machine is also a server.

- **Queue.** The group of customers waiting for service. A backlog of aircraft or components is a queue.

2. Characteristics of Queuing Systems

There are four characteristics to queues:

- **Arrival Process.** The depiction of the timing of customer arrivals at the queue.
- **Departure Process.** The depiction of the timing of customers leaving either the queue or the service.
- **Reneging.** The act of leaving the queue before being served.
- **Service Process.** The time taken to serve customers.

C. HOW QUEUES OPERATE

1. Measurement and Observation

a. Measurement

There are two perspectives in measuring the performance of a queuing system. From the server's perspective, the important measures of performance (MOP) are: service time, proportional utilization (PU), throughput (Global service rate), arrival rate, reneging proportion, and queue length. From the customer's perspective, the important

MOP's are: time in queue, service time, waiting time, proportion of work done on time, and tardiness. [Ref.2:p.20]

PU is important in the ToC philosophy because it measures the torque of the server. PU is the customer's actual arrival rate divided by the maximum continuous service rate of the server. The denominator is the maximum continuous rate because it does not include temporary service rate increases like overtime. As an example, if a server can complete ten customers in a specific time period but only five customers arrive, then the server's PU is .50. With a steady arrival rate, PU is directly and inversely related to the server's maximum continuous service rate. PU ranges from zero to infinity. If PU is zero, there are no arrivals at the server. If PU is between zero and one, the server's maximum continuous service rate is larger than the customer's arrival rate. If PU is one, the server is matching its maximum continuous service rate to the customer's arrival rate. If PU is greater than one, the server's maximum continuous service rate is unable to keep up with the arrival rate of customers, hence queues form (Later the study will show that queues at the NADEP will form with a PU as low as .80). If the server is a piece of machinery with the ability to instantaneously change

its maximum continuous service rate, a PU of one is realistic and ideal. However, in a labor intensive production environment with a union protected fixed work force (such as the NADEP), a PU of one is unrealistic.

A QUEUE at a SERVER is generated when the arrival rate at the SERVER is greater than the SERVER's ability to move customers through. If a QUEUE exists, then the path of the customer is: arrive at the existing QUEUE to await SERVICE, leave the QUEUE into SERVICE, leave the SERVICE. Figure 2 depicts this path. The frequency the customer moves through the SERVER is the SERVICE rate. The SERVICE rate is always equal to the frequency of the customers leaving the QUEUE into SERVICE but not necessarily equal to the arrival rate at the existing QUEUE.

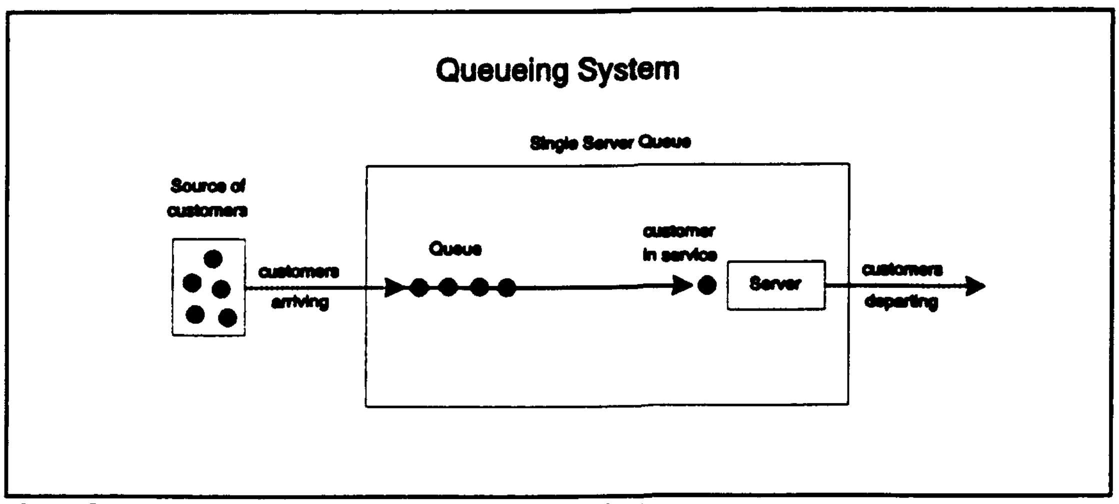


Figure 2: Queuing system with a single queue and single server.

A cumulative arrival diagram shows the accumulation of customers that have arrived over time, and a cumulative departure diagram shows how many customers leave over time. Figure 3 is a cumulative diagram and is important because it provides many of the MOP's in one simple picture. The primary curves of a cumulative diagram are:

$A(t)$ = cumulative arrivals from time 0 to time t

$D_q(t)$ = cumulative departures from the QUEUE from time 0 to time t

$D_s(t)$ = cumulative departures from the SERVICE from time 0 to time t

A representation, in Figure 3, of the number of customers in the QUEUE at any time, t, is equal to the vertical difference between the cumulative arrival curve, $A(t)$, and the cumulative queue departure curve, $D_q(t)$. For example, at time two, there are two customers in the QUEUE. The number of customers in the SERVICE at time t is equal to the vertical difference between the cumulative queue departure curve, $D_q(t)$ and cumulative service departure curve, $D_s(t)$. For example, at time four, there is one customer in SERVICE. The QUEUE or SERVICE TIME for a customer is the horizontal difference between the curves. For example, the QUEUE TIME for customer

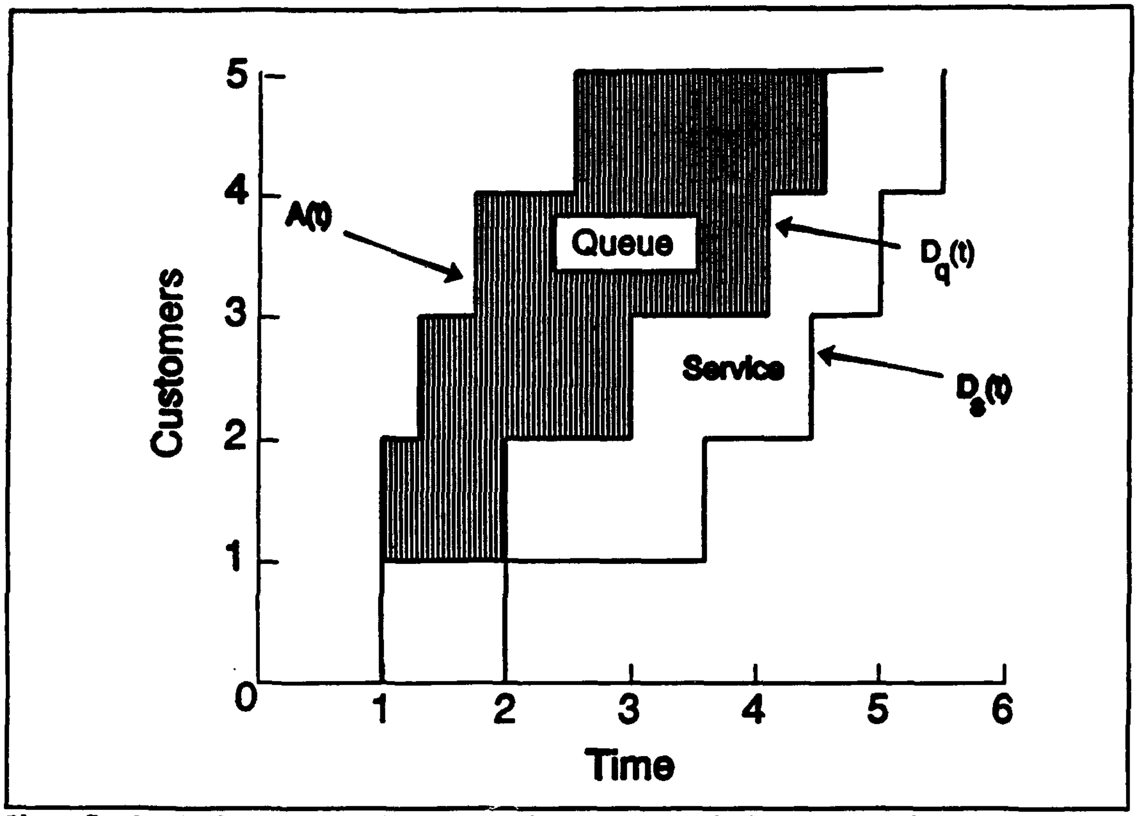


Figure 3: Cumulative arrival and departure diagram. The arrival rate, A(t), is steeper than the service rate represented by $D_s(t)$. $D_q(t)$ is departure rate from the queue.

four is two time units and the SERVICE TIME is one time unit. Time units can be minutes, hours, days, weeks, months, etc.

The ARRIVAL RATE into the QUEUE is the steepness of the arrival curve, A(t). The higher the ARRIVAL RATE the steeper the arrival curve. The ARRIVAL RATE into SERVICE is the steepness of the queue departure curve, $D_q(t)$. The SERVICE RATE of the SERVER is the steepness of the service departure curve, $D_s(T)$, of the SERVER. The higher the SERVICE RATE, the steeper the service departure curve.

b. Observation

The task of measuring cumulative arrivals and departures occur in any of several ways:

- Record the time each customer arrives then leaves from either the queue or service.
- Record the time each customer leaves the service. Periodically count the number of customers in the queue and the number of customers in service.
- Record the time that each customer leaves the service. Periodically time how long a customer spends in queue and spends in service. [Ref.2:p.36]

When measuring queue lengths periodically, the shorter time intervals between measurements enhances the detection of changes.

2. Steady State Queues

In a steady state system, queues materialize because of three factors: exponentially distributed service times, Poisson arrivals, and high PU. These factors are now described.

A typical distribution of remanufacturing service time appears in Figure 4. There are very few service times close to zero, and a zero service time is impossible. As time increases, the frequency of service times increases also. The

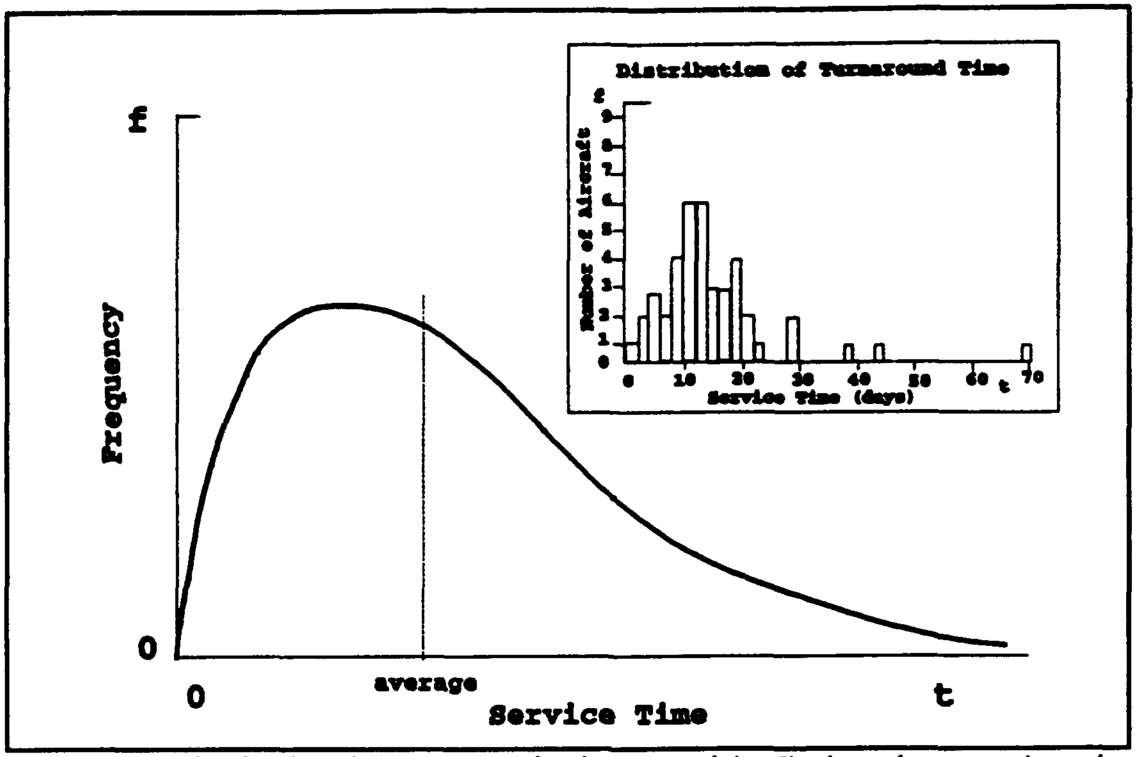


Figure 4: The distribution of the server's time is exponential. The inset is an actual service shop's distribution.

most frequent service time is at the curve's peak. Progressing further on in time, the curve slopes down exponentially. The frequency of service times become smaller but more spread out. This causes the average service time to occur on the right side of the peak. To the far right of the peak, the curve approaches zero asymptotically because, theoretically, a customer could be in process indefinitely. The inset in Figure 4 is the actual distribution of service time of one shop (Shop 23) on the A-6 remanufacturing line. The exponential distribution is clearly seen.

Poisson arrivals meet the following criteria:

- The probability that a customer arrives at any time does not depend on when other customers arrive. For example, aircraft do not arrive at exactly 23 day intervals.

- The probability that a customer arrives at any time does not depend on any specific time interval. For example, aircraft do not arrive at the depot on only the first of the month.

- The customers arrive in quantities of one.

PU was previously described as equal to the customer's actual arrival rate divided by the server's maximum continuous service rate. If a system of servers are steady state but experience a Poisson arrival pattern and exponentially distributed service times (like the *NADEP*), a PU of greater than .80 causes queues to be significant. Figure 5 presents a representation of the average customers in a queue versus PU. No significant steady state queues exist below a PU of .80.

3. Non-stationary Arrivals and Fluid Approximations

A non-stationary arrival pattern occurs when the customer arrival rate varies over time. From a server's perspective, it is far easier to serve customers when they arrive at an even rate than at an uneven rate. In the

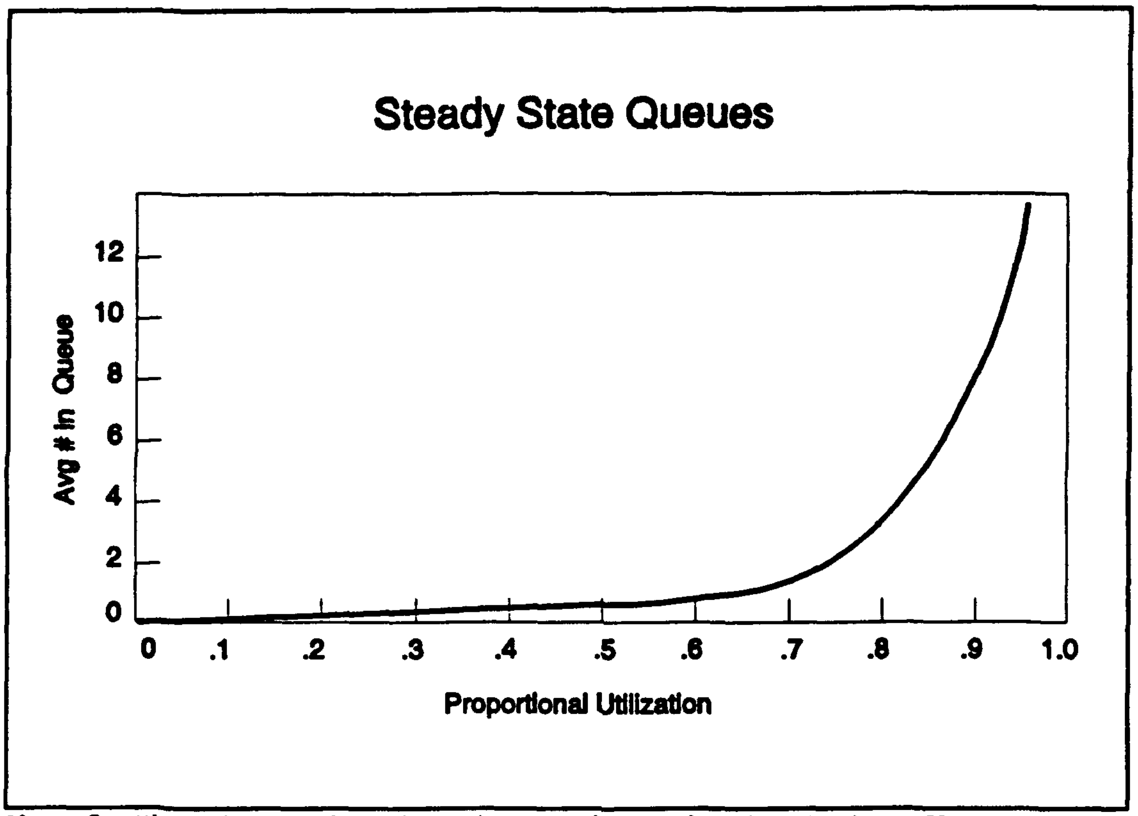


Figure 5: With Poisson arrival time and exponential service time, PU above .80 causes average queue lengths to grow. [Ref.3:p.15]

previous section, steady state queues formed at random because of variability in arrival times, service times, and because of high PU. A non-stationary arrival pattern presents an additional source of variability that is large but predictable. The arrival pattern in this system is predictable because the queues occur in a predictable manner. Since the variability is large, the servers have a difficult time managing the arrival of customers, hence large queues form. The larger the queue, the more costly it is to manage. Rush hour traffic is a good analogy of predictable variability.

Rush hour traffic is a good analogy of predictable variability.

Figure 6 presents predictable arrivals with minor perturbations reflecting random variation. The larger the number of arrivals, the smaller will be the effects of the random variations and the more deterministic the arrival process. A fluid model can approximate customers with a large induction or arrival rate. A fluid model is a very robust tool to use when analyzing a queuing system. In a fluid model, customers are viewed as a liquid that accumulates in a

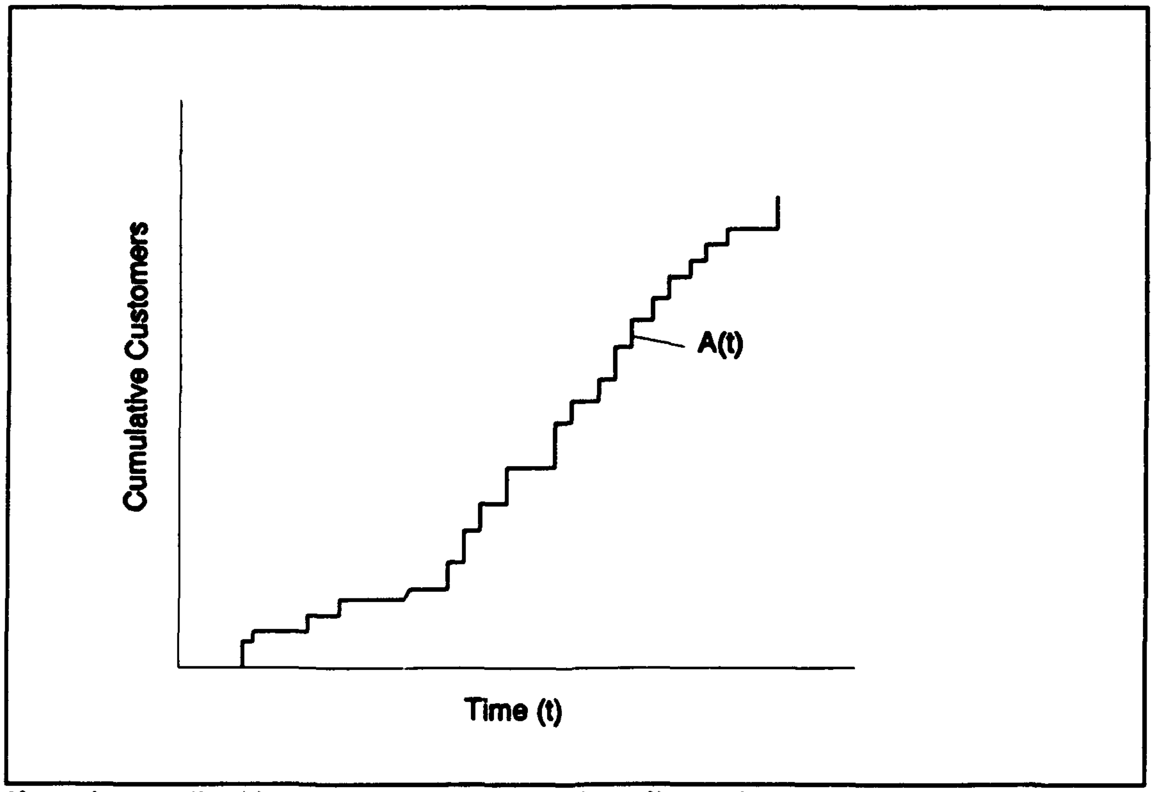


Figure 6: predictable arrival pattern with minor fluctuations over time. The pattern is predictable because the slope of the line is predictable.

tub. Queues increase when the fluid enters the tub faster than it leaves and queues decrease when fluid leaves the tub faster than it enters.

The non-stationary arrival pattern (arrival rate varies over time) divides into four distinct phases. Figure 7 portrays the non-stationary arrival pattern:

1. Phase one, Stagnant, the customers are being served as fast as they arrive. The queue is either empty or steady state.

2. Phase two, Queue Growth, the customers arrive at a faster rate than the service rate so the queue grows.

3. Phase three, Queue Decline, the customers arrive at a slower rate than the service rate so the queue shrinks. Notice from Figure 7 that a queue can exist even if the arrival rate is smaller than the service rate.

4. Phase four, Stagnant, queue vanishes because the customers' arrival rate continues to be less than the server's service rate.

In all four phases, the arrival rate and the service rate determine when the queue grows or shrinks. The queue is largest at the end of phase two because this is the last time the decreasing arrival rate equals or exceeds the steady service rate. This is also the time the queue begins to shrink. The queue disappears when the arrival rate is much

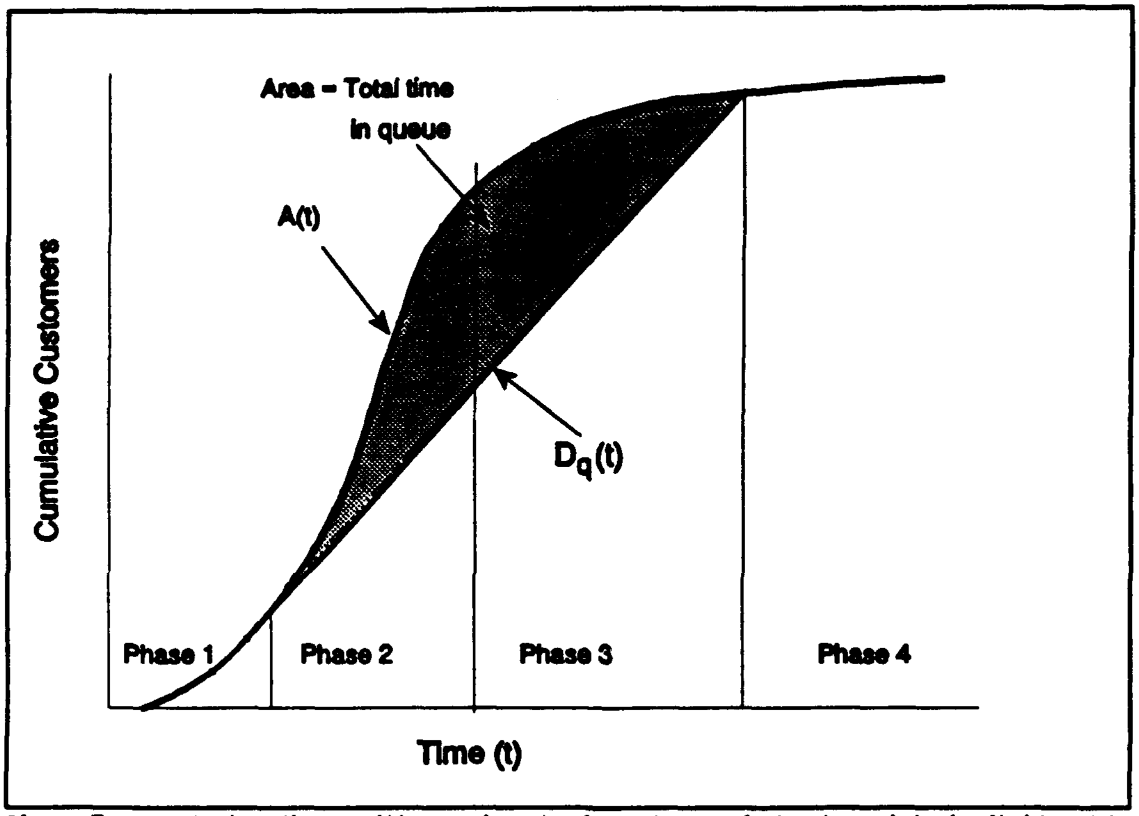


Figure 7: Cumulative diagram illustrating the four phases of the deterministic fluid model: stagnant, queue growth, queue decline, and stagnant.

smaller than the service rate. The length of time the queue takes to disappear depends on the difference between the service and arrival rate. Unless the service/arrival rate differential is large, a service system with a large queue and a large service time will take a long time for its queue to vanish. A large induction queue at the NADEP caused by a period of high arrival rate, could take several years to vanish. The NADEP, Alameda, CA, A-6 remanufacturing line is an example of this problem. When the queue at the server

(NADEP) has finally vanished, it will be critical to control the arrival rate to not allow the queue to grow again.

The deterministic fluid model is the best way to visualize a non-stationary queuing system. This fluid model does not account for random fluctuations in the arrival and service processes, just predictable fluctuations. These predictable fluctuations overwhelm smaller, random fluctuations in arrival times.

Figure 8 illustrates the effect on turnaround time (TAT) because of non-stationary queues. The departure rate ($D_q(t)$) from queue (also equal to service rate) is steady state. The arrival rate, A(t), in Figure 8(a) increases predictably and once past the $D_q(t)$, the queue length in Figure 8(b) grows. When the arrival rate in Figure 8(a) falls below $D_q(t)$ (slope of A(t) is less than the slope of $D_q(t)$) the queue length is maximum (note time a in Figure 8(b)). As the queue increases, the TAT in Figure 8(c) also increases but at a slower rate. When the queue is at its maximum, the TAT is increasing at the steepest rate. The TAT does not decrease until the queue has reduced substantially (note time b). This time lag effect causes TAT to be a poor predictor of existing work backlogs. To decrease TAT in the remanufacturing environment, reductions

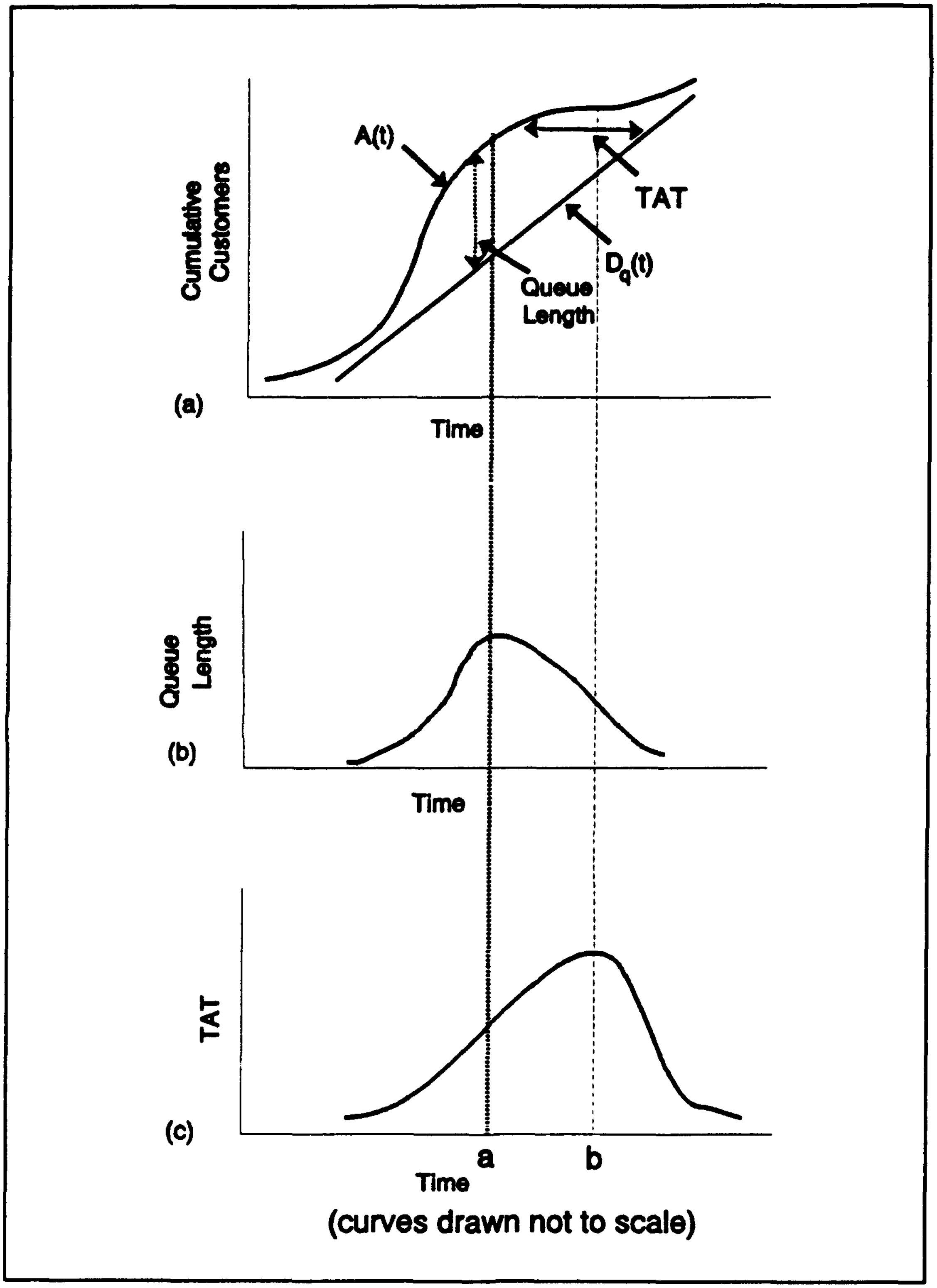


Figure 8: The effect of queue growth on TAT. Figure (b) and (c) come from (a).

in queue length and WIP must be maintained to allow the lag to work itself out. Many efforts to reduce TAT by reducing WIP have been abandoned too soon. Also, using TAT to measure the quantity of work to be inducted into service is a bad idea because of the lag effect shown in Figure 8. The lag time is the measured time difference between a and b on curve (c).

D. REDUCING QUEUE TIME

Queues are real quality costing problems. The manager's sensitivity to keeping the queues at an optimal level is critical to maximizing productivity.

1. Queuing Problems

Queuing problems are classified into three types:

- Perpetual Queue. All customers have to wait because the servers do not have the ability to handle the demand for the service.
- Predictable Queue. Previously discussed.
- Stochastic Queue. Customers arrive by chance faster than the average service time. [Ref.2:p.208]

Each of these queues needs special attention. The perpetual queue is a sign that the server disregards the customers' time and only seeks to minimize the cost of

providing the service. Many of the queues at the NADEP are in this category. The solution is simple: do something to either increase the service rate or decrease the arrival rate. Elimination of a predictable queue occurs by increasing the service rate during "peak hours". Stochastic queues are the most difficult to reduce because they are not predictable. Fortunately, stochastic queues are small and short lived so it is not always critical to reduce them.

The two types of servers in queuing problems are service personnel and specialized equipment. The strategy to reduce queue size is different for each.

2. Match Service Process With Arrival Process

The most obvious reaction to decreasing a queue formed by an increasing arrival rate is to add another server at a predetermined critical queue level. This works very well in a grocery store check out line but when the server is a union worker with prohibitions in job rotation or if the server is an expensive specialized piece of equipment, this will not work. In these cases, increasing the average service rate and emphasizing reducing service time is one method of decreasing queue size. Another method to reduce queue size is

selectively increasing the service rate at times when queues exist.

a. Decrease Service Time and Increase Service Rate

Servers with queues do not have the problem of customers arriving into service with a Poisson arrival pattern. That occurs at the entrance to the queue and is neutralized by the queue itself. The arrivals into service will occur at the rate the customers depart the service. A server with this condition will have a PU of one. It is important this server has the highest priority for resources that will increase its maximum continuous service rate. At the NADEP, this means adding extra artisans on a long term basis to the resource priority service shop. When adding the extra resources, the initial emphasis should be to reduce service time rather than to increase service rate. Figure 9 illustrates that an increase in service rate both decreases the size of the queue and the duration of the queue. This means the relationship between total waiting time (the area between the arrival and departure curve) and the service rate is nonlinear. When queues are large, SMALL changes in service rate create large changes in waiting time. A 25 percent

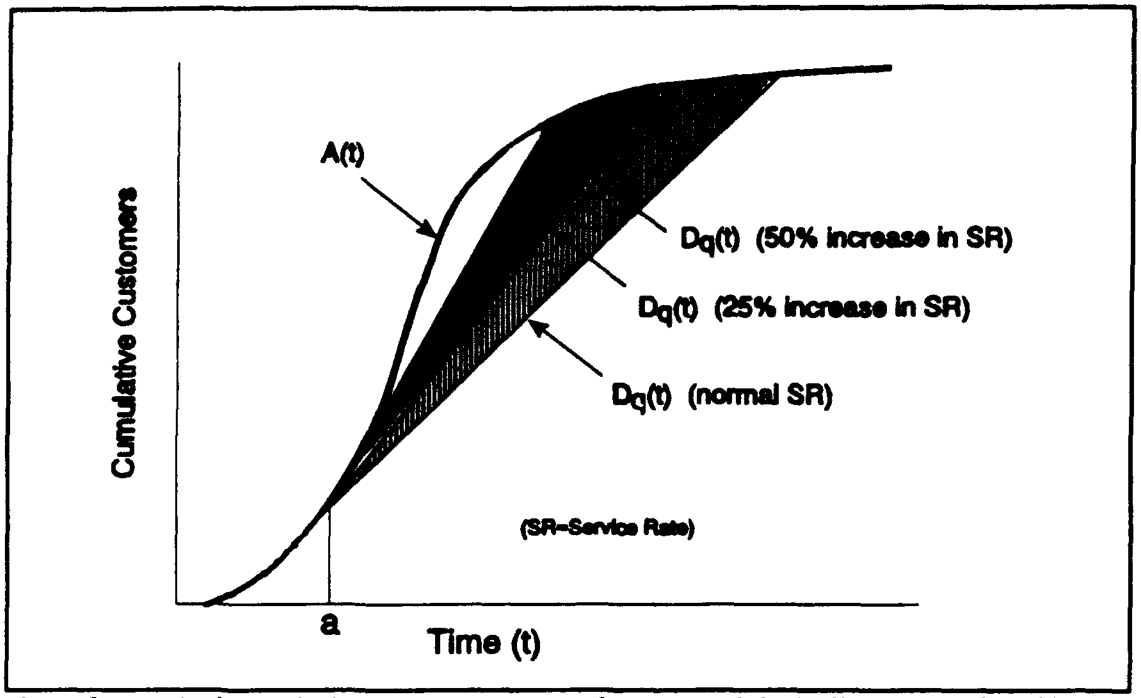


Figure 9: Cumulative arrival and departure curves for a deterministic fluid model with different service rates. A small increase in service rate at time (a) creates a large decrease in waiting time.

increase in the service rate in the Figure 9 example produced roughly a 50 percent reduction in the waiting time.

b. Varying The Service Rate

Figure 10 shows a hypothetical cumulative arrival curve (A(t)), normal service departure curve ($D_s(t)$), and three parallel, supplemental service curves (A, B and C). The supplemental service curves are parallel because their service rates are the same. The difference in the three supplemental service curves is the scheduled time they are turned on. Curve A is turned on in sufficient time to cause the queue to disappear temporarily (note the tangent point). The shaded

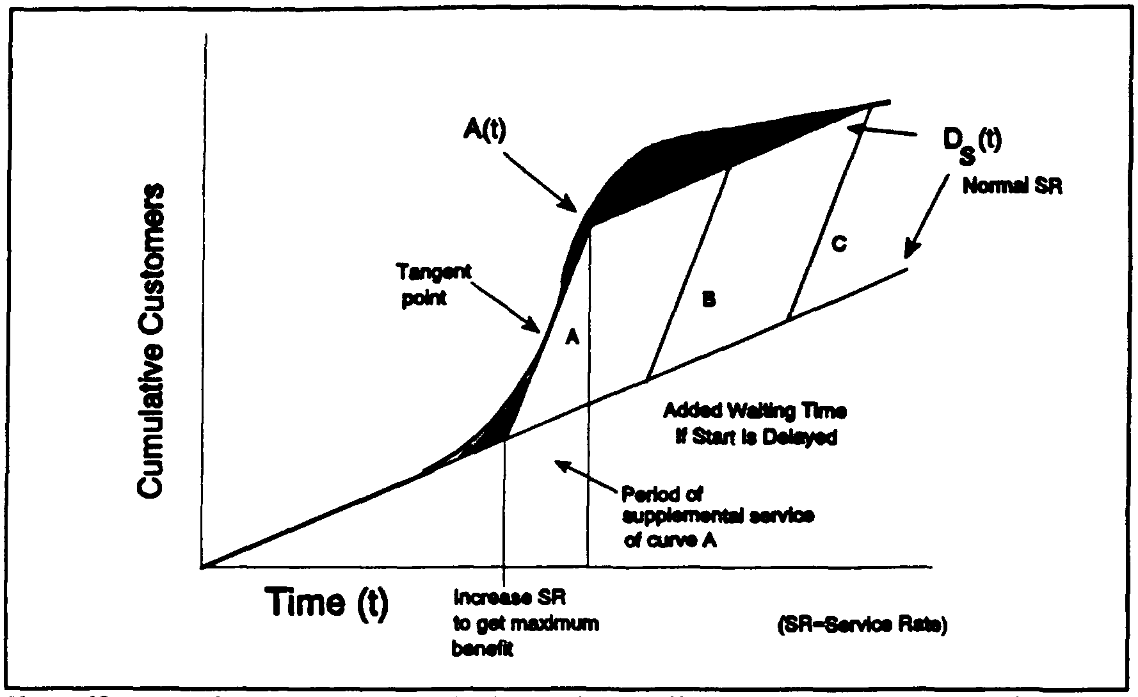


Figure 10: A service rate supplement should be added sufficiently early so the queue will vanish at least once during the supplemental period.

area represents the remaining queue after using supplemental service curve A. Curve B and C are turned on at later times with the effect of only partially reducing the queue. The following guidelines help when adding supplemental service rate:

- Service should be in place before the queue is formed (supplemental service includes artisans, tooling, hangar space, training, etc).
- The increase should occur early in the queue (curve A in Figure 10), to ensure that the queue vanishes sometime before the added service is removed.
- The duration of a supplemental period depends on a trade-off between service costs and waiting costs, as does the size of the service supplement.

- High service cost is an incentive for reducing the duration and size of the supplemental service.

If both the number of servers and the duration of the supplemental period are flexible, then service rate varies continuously, exactly matching the arrival rate. The ideal situation is maximum PU with zero waiting time. A footnote to the study is that Figure 10 can also model the delay caused if the server breaks down or takes a break. The departure curve, $D_s(t)$, drops off, increasing the queue lengths. A model of an anticipated service rate drop off (seasonal holiday) and its effects on TAT and production flow allows for future planning of the arrival rate of customers.

3. Match Arrival Process With Service Process

When the arrival process more closely matches the service process, the total time spent by customers waiting reduces.

The queue developed in Figure 10, because the arrivals, $A(t)$, exceeded the service rate, $D_s(t)$. If the server can not increase its service capacity then it should require each customer to arrive at the rate it would normally leave the queue, that is, $A(t) = D_q(t)$. This causes the queue to become

steady state. The customer still exits service at the same time as before but now he has a set arrival time or reservation depending on the forecasted length of service time. A reservation system nullifies the effect of Poisson arrivals. The contribution of reservations is the removal of predictable queues. Reservation systems can also help eliminate random queues by reducing variation in customer arrival times. When using reservation systems, the constraint servers should operate with a PU less than 1 to allow for a surge with anticipated late arrivals

E. QUEUING NETWORKS

The majority of service processes, other than grocery store check outs or bank tellers, are multi-servers. A multi-server is a process with more than one server. Each additional server may or may not be dependent to the other servers in the process. The most important concept in multi-server queuing networks is the constraint. In most queuing networks, the performance of one or more of the constrained servers influence queuing delays. To improve the performance of the total system, the constraints are first identified, then their performance improved. Constraints are very

important in each of the two types of network service, serial and parallel. With serial service, Figure 11(a), the server accepts one consecutive customer at a time. Notice that server one's departure process dictates server two's arrival process. With parallel service, Figure 11(b), servers perform tasks simultaneously. Many networks can contain a mixture of both, working interactively, to produce components for the

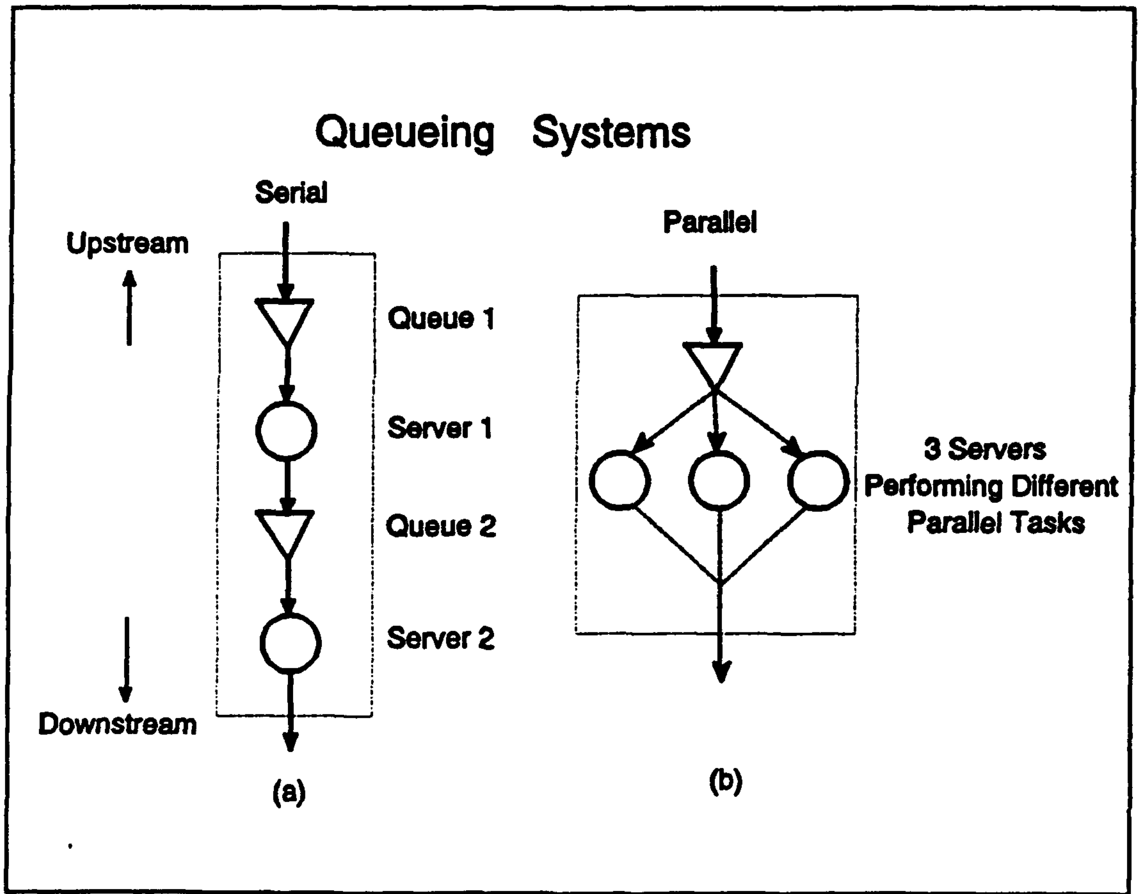


Figure 11: Diagram (a) illustrates two servers in series showing, upstream and downstream directions. Diagram (b) illustrates three servers in parallel, fed by a single queue.

production of a single product. Developing analytical models of queuing networks is extremely difficult because of their stochastic nature. One way to understand queuing networks is with graphical models. The most important system measurement of queuing networks that is measurable with graphical models is the rate of output.

1. Fluid Model of Series Queue

If in a multi-server system, service to customers is continuous (no batch processing), has short service times compared to the total expected time in process, and has an arrival rate that varies deterministically over time, then the system can be analyzed with a fluid model. Figure 12 represents a graphic fluid model of a server process. The graphs are very robust in information concerning the performance of the server system. The following are specific definitions concerning the graphs:

- **Time in system.** This is the difference horizontally between the arrival time, $A_q^1(t)$, at induction into the system and the departure time, $D_s^n(t)$, from the last server, server n.
- **Customers in System.** The total number of customers in the network is the difference vertically between the arrival curve, $A_q^1(t)$, and the departure curve, $D_s^n(t)$.
- **Time in Service.** The sum of all the service times.

- **Time in queue.** The difference between time in system and time in service.

In Figure 12, both servers have queues in front of them but only one is the real constraint. It would appear that server one is the constraint because it has the largest queue, however it is actually server two because it has the smallest service rate, $D_s^2(t)$. With queues in series, the server with the smallest service rate will also be the last server to have

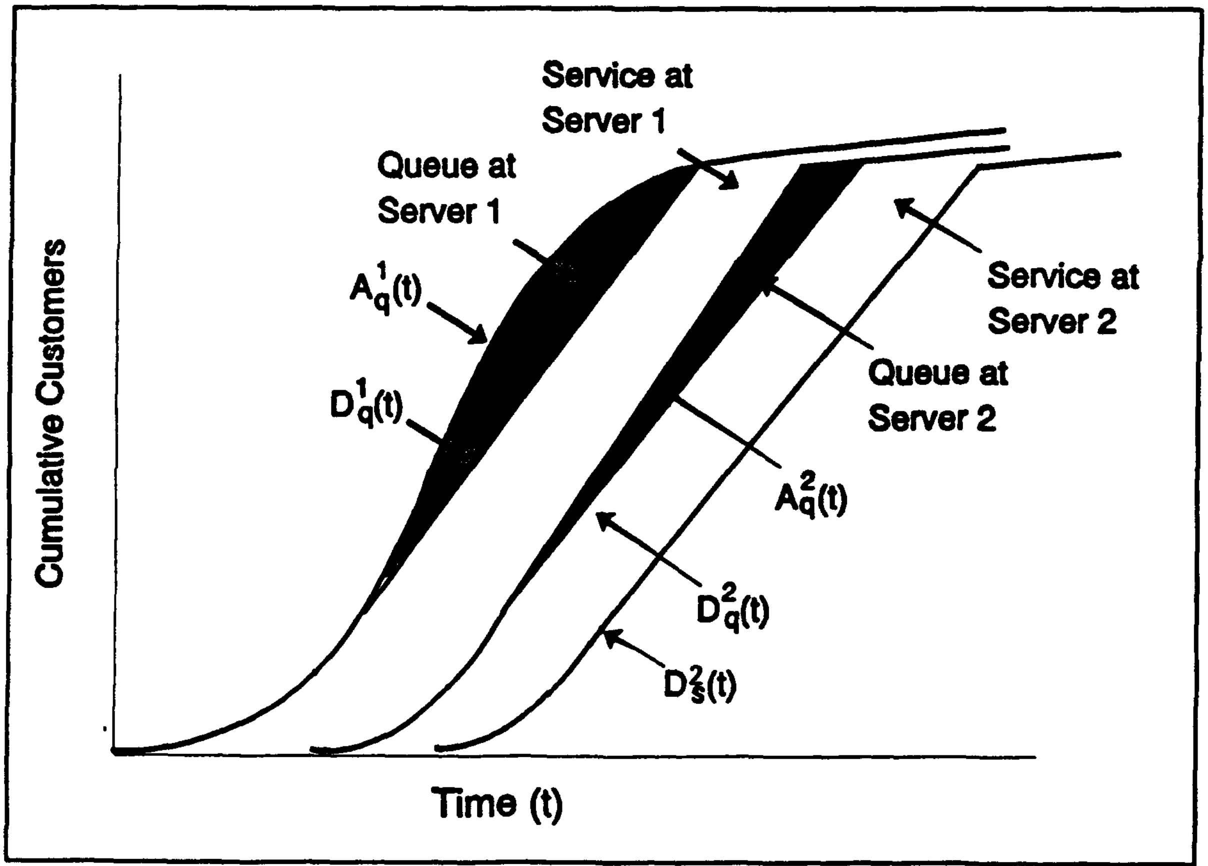


Figure 12: Cumulative arrivals and departures for a deterministic fluid model of two servers in series.

a queue. This server is the actual constraint. The service rate of the constraint server directly affects the time customers spend in a series queue. If the series queue contains more than one constraint, then all the constraint service rates should be balanced. Increasing the service rate of non-constraint servers will not decrease total time in the system. The time in service may decrease but time in queue will increase. So, to increase flow, the process cumulative arrival and departure queuing charts make it very easy to find constraints, even in a complex remanufacturing environment. In the next chapter, we will see how to use these charts to apply ToC to aircraft remanufacturing.

2. Fluid Model of Parallel Queues

Figure 11(b) depicted the simplest parallel server diagram. A job is not complete until each server completes its task. Queuing models containing both serial and parallel queues are more representative of the remanufacturing process. Figure 13 represents such a model. A customer enters the system, gets broken down into components, is serviced, then is reassembled before exiting. Two dimensional fluid models work for serial servers only. To model parallel servers, either a

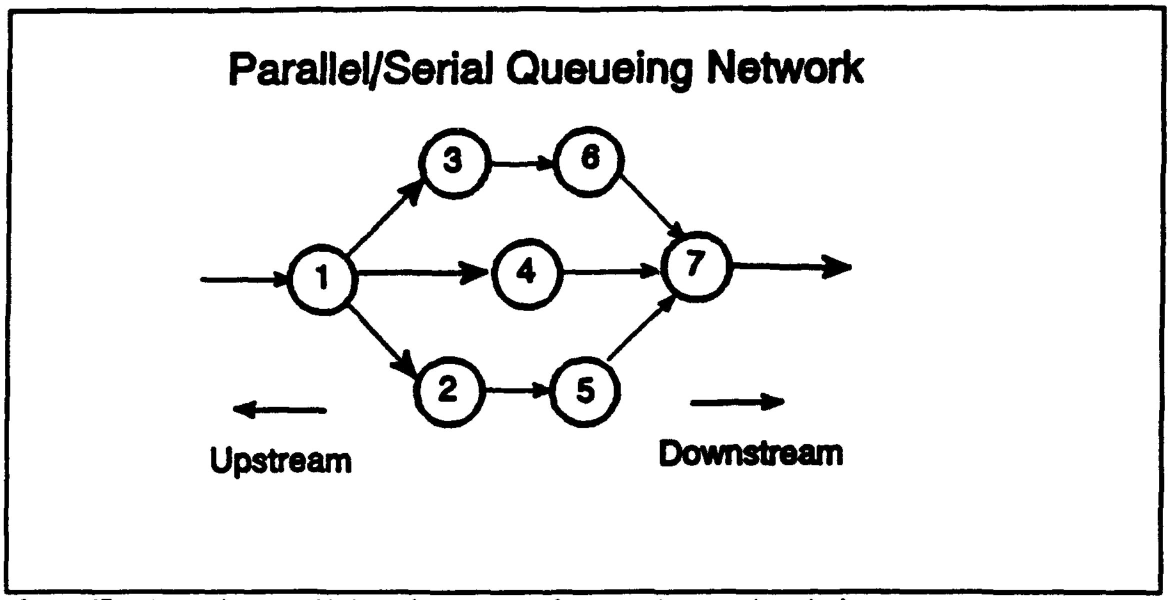


Figure 13: A complex parallel/serial server diagram, where each node is a server.

three dimensional fluid model is needed or each parallel line of servers is individually modeled. Identification of the constraints is difficult with the latter method.

3. Queue Discipline in the Network Model

Queue discipline should be global vice local. Global discipline is achieved by first determining the constraint through already presented techniques, then making the constraint the priority.

Constraints should never be idle for lack of work unless the market is the constraint. A buffer of customers at the constraint cushions it against random variation in arrival and service rates. However, not all customers visit constraints,

and those that do not should not be sitting in queues awaiting service from a non-constraint.

4. Induction Queues and Bulk Service

When there is a large variation in the induction, the first server requires a smoothing queue, or buffer. The smoothing queue will prevent the fluctuations in the inductions from being passed on.

Bulk service is worth noting here. If customers come to the servers in bulk (in large batches), the number of setups reduce. However, reducing setups increases the queue to downstream servers. This may outweigh the benefit of reducing the number of setups.

F. CONCLUSION

The methods of observing and measuring a queue are important if an analysis is undertaken. Queuing, whether at the local fast food restaurant or an aircraft remanufacturing line, breaks down into two phases, the arrivals to the process and the inclusive system. Each phase has unique characteristics that affect each other on a total system performance level. Models enhance the analysis of a server process for queuing problems. One particular model, fluid

approximation for non-stationary arrivals, is a particularly good model because it filters out random variation in arrival times and looks purely at the predictable arrival rates. Finally, solutions were introduced as ways to begin reducing costly delays in a server process.

IV. THEORY OF CONSTRAINTS AND NADEP REMANUFACTURING

A. INTRODUCTION

The purpose of this chapter is to explain relevant topics from Theory of Constraints (ToC) that apply to aircraft remanufacturing. Familiarity with reference four, The Goal, and reference five, The Race, or Just-In-Time production literature would facilitate the understanding of this chapter.

B. THEORY OF CONSTRAINTS

A constraint is "anything that limits the organization from achieving higher performance versus its goal." [Ref.5:p.106] ToC is a technique for operating a depot using a form of management by exception. It focuses on the areas that have a dramatic impact on the whole depot.

There are two objectives of the traditional approach to managing a depot. The first objective is to focus attention on trying to maximize the efficiency of each server. The second is to use customers (inventory) throughout the shop floor to be sure that all servers are operating with high efficiency and constantly producing. [Ref.4:p.84] The ToC

approach is to place a buffer, or safety stock, of customers in front of servers that constrain the system and no where else. The purpose of the buffer is to be sure that a constrained server is not shut down for lack of customers (The proportional utilization (PU) of the constrained server should never drop below a maximum of one). Because the Naval Aviation Depot (NADEP) has a fixed, union protected workforce, the non-constrained servers at the NADEP will have low PU. That is, they will continue to maintain the same maximum continuous service rate capability but experience a low (or zero) arrival rate. The non-constrained server then will sit idle when there is not a requirement for them to be active. Scheduled idle time is a tough concept for the traditional manager that believes every server must be always 100% active.

1. Basic Concepts

a. Dependent Events and Statistical Fluctuations

Two characteristics compound problems of planning and control of NADEP remanufacturing operations: dependent events and statistical fluctuations. [Ref.6:p.658] Dependent events exists when sequence operations have to occur to produce a part. The process routings that dictate the order

in which service operations are to occur are dependent events. Statistical fluctuations are the day-to-day differences in servers' output causing a random nature in the completion time of a customer. Together, these two phenomenon cause constraints to the operations of the depot.

b. Measuring the Process

ToC measurements are:

- **Throughput.** The rate at which the system generates money through sales. This is equal to total production sales minus material purchased in manufacturing.

- **Inventory.** All the money invested in materials bought for manufacturing.

- **Operating Expense.** All the money spent to turn inventory into throughput.

c. ToC's Five Steps

Two types of servers exist:

- **Constraint.** Any server that is operating with a PU of one but still restricts throughput.

- **Non-constraint.** Any server not operating at high PU (PU<<1). Operating a non-constraint at high PU will only increase work in process (WIP), but not throughput.

ToC emphasizes management of the constraint servers. There are five steps according to Goldratt to help focus on the specific physical constraint:

1. Identify the constraint.

2. Decide how to exploit the constraint.

3. Subordinate everything else to the above decision.

4. Elevate the constraint.

5. If in the previous steps, the constraint is broken, go back to step 1 but do not let inertia become the constraint. [Ref.7:p.2]

Managing by this five step process is ongoing. ToC manages constraints and non-constraints by using an approach called Drum-Buffer-Rope.

2. **Drum-Buffer-Rope**

The Drum-Buffer-Rope (DBR) concept of Goldratt is the scheduling technique of ToC. Ideally, all non-constraints, before the constraint, should begin work on the customer (aircraft or component) as soon as it arrives at the first station in its routing. Thus, the customers move very quickly from material release to the constraint buffer. The customers then wait until the constraint begins to process them.

Chapter V discusses Critical Resource Scheduling (CRS) and how it requires the customers to arrive at the constraint at a certain time to act as a time buffer vice a material buffer. Ideally all non-constraints on the routing between the constraint and the final downstream server should also set up for the customer. The customer should then move one unit at a time to completion. The DBR schedule development consists of three stages:

1. Develop a detailed schedule for the constraint. This schedule, called the drum, also communicates a PU level for each service shop.

2. Determine the time allowed for a customer (aircraft) to move from release to the constraint. This time offset, called the rope, links customer arrivals (aircraft inductions) to the constraint. The constraint pulls forward the customers as it needs them.

3. Determine the total turnaround time (TAT) left to completion when the customer leaves the service of the last constraint. This accurately forecasts a customer completion time.

C. SUMMARY

Theory of Constraints approach to remanufacturing planning and control system design is to accept the existence of an unbalanced process. The constraint is the most limited server. ToC breaks dependencies by creating buffers but only

to buffer the constraint. Non-constraints are already operating with reduced PU and hence have excess service rate. ToC states that buffering them with excess customers (inventory) is a waste. However, the constraint is operating at a high PU, buffering it only adds intrinsic value and hence is not a waste.

V. SYNCHRONIZING AN AIRCRAFT REMANUFACTURING LINE

A. INTRODUCTION

In this chapter, Theory of Constraints (ToC) as applied to the A-6 line is described first, then followed by a discussion of the queuing charts and ToC's Drum-Buffer-Rope (DRB) strategy in relation to the A-6 line. Finally, Critical Resource Scheduling (CRS) is presented.

B. MEASURING THE REMANUFACTURING PROCESS

ToC measurements as applied to the A-6 line are:

- **Throughput.** The rate at which the depot generates money through aircraft completions. It is the money paid to the depot for each reworked aircraft, minus material purchased to rework that aircraft. The shop service rate and proportional utilization rate is in direct relation to the A-6 line throughput rate.

- **Inventory.** The money invested in the material bought to rework the aircraft that the depot intends to deliver to the fleet is inventory. The aircraft and components are also considered inventory. The inventory level is directly related to the level of inducted aircraft.

- **Operating Expense.** The money spent to support the aircraft rework process. This expense includes all salaries, wear on tools, depreciation expenses, heating, lights, etc. incurred in the remanufacturing process.

C. QUEUING CHARTS

The three basic elements of the queuing system are customer, server, and queue. The customer at the A-6 line is a complete aircraft at induction but immediately becomes component parts and a structural hull at disassembly, Shop 21. This shop, the rest of the A-6 shops, and the component rework divisions are the servers of the multi-server queuing network that makes up the A-6 line. Each server is in series, in parallel, or a combination of both. The component parts and the hull structures that are not being worked on for lack of labor or parts are the queue. Figure 14 is a fluid queuing chart depicting inductions and completions of all A-6 aircraft for FY 89-2 to FY 92-2. The horizontal distance between the two curves is the average aircraft turnaround time (TAT). The vertical distance between the two curves is the number of aircraft-in-process (AIP) and relates directly to inventory. The negative numbers on the vertical ordinate (cumulative aircraft) represent aircraft that have completed the process. The positive numbers are aircraft still in process. The slope of the induction curve, $A(t)$, is the arrival rate. The slope of the completed curve, $D_s(t)$, is the departure rate. Throughput is the slope of $D_s(t)$ (translated into sales

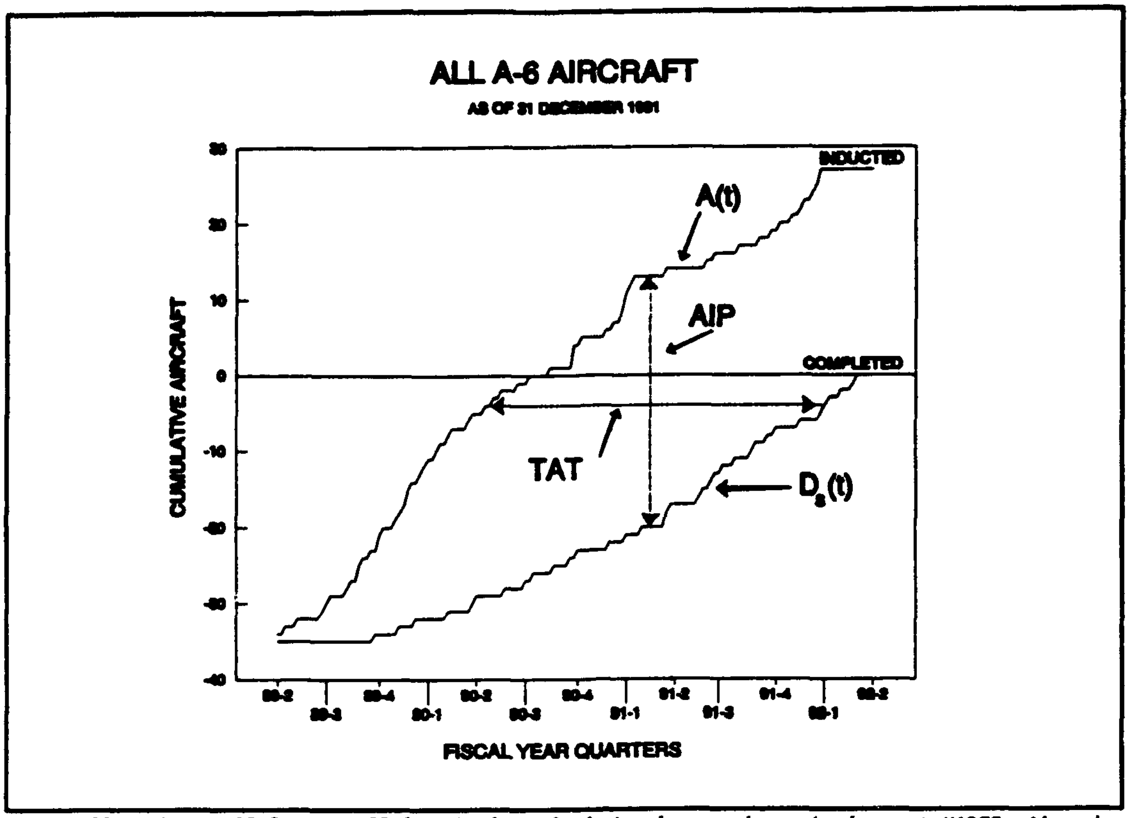


Figure 14: The FY 89-2 to FY 92-2 A-6 aircraft inductions and completions at NADEP, Alameda. Horizontal distance is average TAT and vertical distance is AIP.

dollars) minus the cost of materials during the same period used to measure the slope of $D_s(t)$.

1. Analysis of A-6 Data

Starting in FY 89-4, AIP increased due to the high arrival rate of rewing aircraft. The A-6 line had a steady completion curve until FY 91-2. In FY 91-2, the effect of FY 91-1 Department of Defense (DOD) civilian wide hiring freeze was being felt. This freeze, still in effect, has prevented the hiring of additional workers. A-6 aircraft completions

increased after FY 91-2 for two reasons. One was the decision to move the majority of P-3 aircraft work to NADEP Jacksonville. This allowed 97 artisans to transfer from the Alameda P-3 line to the Alameda A-6 line. The second was the approval of a waiver to hire an additional 33 artisans. As a result, A-6 aircraft completions increased after FY 91-2 as total artisans assigned to the A-6 line rose to over 330. Unfortunately, these delays in adding supplemental service added considerably to the average TAT (see Figure 9 and Figure 10, Chapter III). Also, an attrition rate of about 30 employees a month from NADEP Alameda (with the A-6's share at two to three a month) steadily eroded the gains from the supplemental service.

The A-6 line appears to be at the end of phase II, queue growth (see Figure 7, Chapter III). Average TAT will remain high until AIP returns to its pre-FY 89-4 levels (Figure 8, Chapter III).

2. Fluid Approximation Plot of I Logical Structure

Applying the idea of accumulating system induction and completion rates, then plotting them with each other to develop the cumulative curves, works with the individual shops

as well. If the process is taken one step further and all the plots for the shops are combined, the fluid approximation queuing chart representing the complete A-6 line is produced. Figure 15 is the fluid queuing chart for the A-6 line.

Each curve in Figure 15 is the induction curve into the next shop but also the departure curve for the previous shop. Again, the horizontal distance between curves is TAT. The vertical distance between curves is AIP.

3. A-6 Line Fluid Approximation Curves

The shop curves in Figure 15 began to diverge in the FY 89-4 time frame. Before that, the shop curves appear parallel and steady state. Although not depicted, since before FY 89-2, the departure curves for shops 31 through aircraft flight test are parallel to the final departure curve. However shops 21 through 26 show large variation in TAT and AIP. The Acceptance, Strip, Disassembly, and Evaluation and Examination shops had to cope with a large variation in arrivals. They were able to handle the variation in arrivals because they were operating with low proportional utilization (PU). There was not a buffering for the variation in arrivals so the variation in AIP and TAT continued through

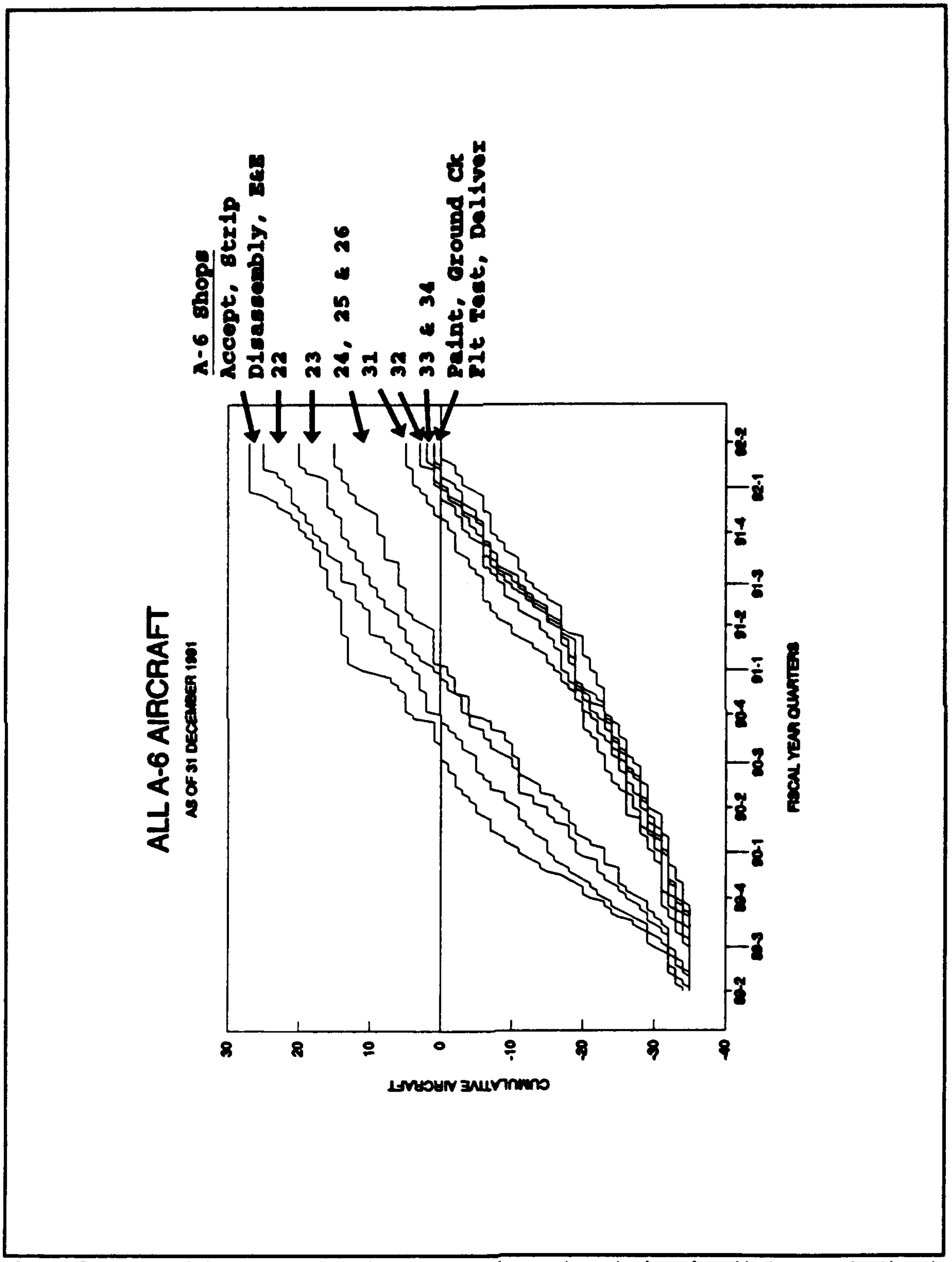


Figure 15: The FY 89-2 to FY 92-2 A-6 aircraft inductions and completions for all shops on the Alameda A-6 line. Note the increasing TAT and AIP in shop 24, 25 & 26.

the rest of the structure's shops, 21, 22, 23, 24, 25, and 26. The arrival process was not predictable, especially at FY 90-4 where the effects of batch inductions are seen in the downstream shops.

The shops, 24, 25A, 25B and 26, are parallel/serial shops as shown in Figure 1, Chapter II. Shop 24 is SDLM rewing. Shop 25A and 25B are the composite rewinging process shops which began in FY 89-4. At this same time, the arrival rates grew and the queues first began to grow at the depot. Figure 16 shows individual shop queuing charts.

4. Location of Drum, Buffer, and Rope on Queuing Charts

The rewing process is the constraint because it is the last place on Figure 15 where a queue forms. Figure 16 shows that Shop 25A, composite wing installation, has the longest TAT and AIP. Much of this time was delays in receipt of the composite wing from Boeing Aircraft (see Figure 17) and as a result the AIP grew. The production rate of Shop 25A is the Drum Beat. Shop 25A requires a Buffer to ensure the PU is maintained at one, ie the Drum does not miss a beat. The next section will show the drum beat's buffer, through better managed resource scheduling, is a time buffer rather than a

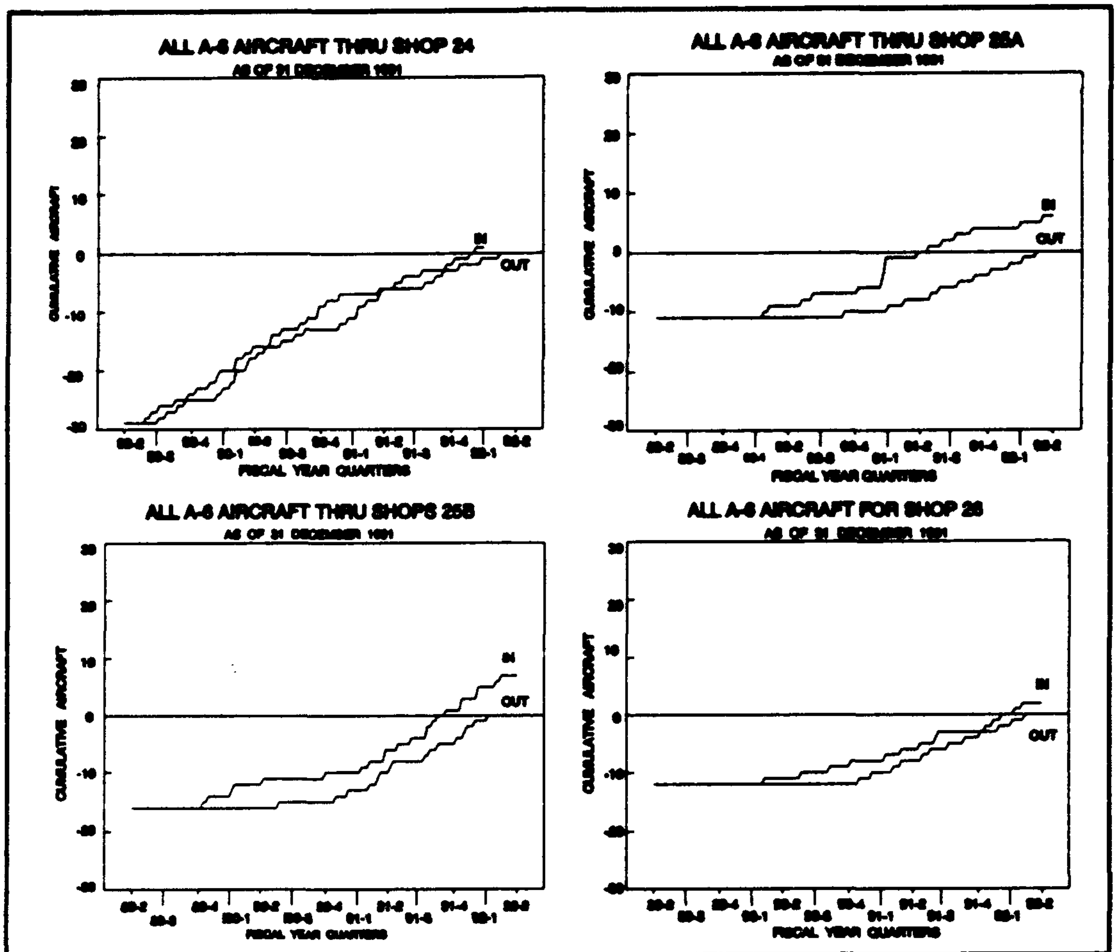


Figure 16: Shops, 24, 25A, 25B and 26, induction and completion curves are shown. Shop 25A and 25B have the only queues. Shop 25A is the constraint because it has the last queue and smallest service rate.

material queue buffer. The Rope should be tied from the constraint, Shop 25A, to NAVAIR to synchronize the arrival rate of aircraft with rewing production.

Figure 17 illustrates production flow of the four A-6 lines with the use of queuing charts. The four queuing charts show induction and completion curves for the aircraft as they arrive and leave the A-6 line. The SDLM A-6, in Figure 17(a),

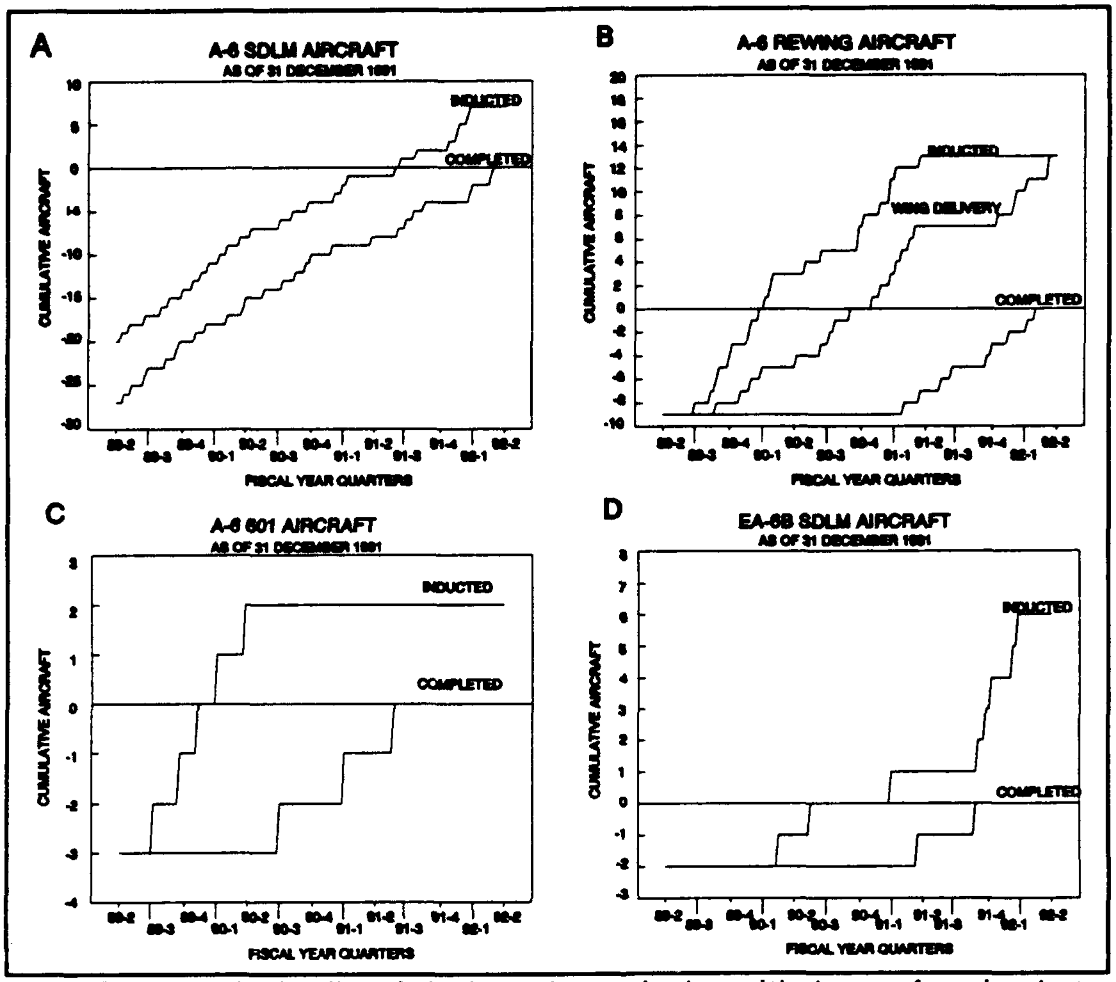


Figure 17: The production flow of the four A-6 types is shown with the use of queuing charts. Note the divergence in the curves (growing queue) of the composite rewing operation.2

experienced a steady state flow because the induction and completion curves remain parallel. The 601 A-6, Figure 17(c), have divergent curves representing the difficulty the depot had in implementing the 601 wing structure modification. The curves even depict the failure of 601 as an airframe change because two aircraft have been in process since FY 90-1. Figure 16(d) illustrates the SDLM EA-6B flow. The first two

aircraft took about five quarters to complete the process. Five EA-6B's were inducted at a high rate in FY 91-3. They are currently behind schedule and experiencing high TAT. The induction rate of these five aircraft is an example of batching. The A-6 composite rewing flow curves are shown in Figure 17(b). The curves are highly divergent representing the high arrival rate of aircraft and the lack of service capacity. Note on Figure 17(b) that the receipt of the composite wing structure occurs at a slower rate than the arrival of the A-6 aircraft. Since the rewing shop is the constraint, either the arrival of the wing structure should be earlier or the aircraft arrival should be later. Supplemental resources (artisans, hangar space) should be allocated to the composite rewing shop to increase the service capacity, therefore increasing flow through the constraint. It is also critical this constraint server is scheduled to obtain maximum performance (see p. 32, Varying The Service Rate).

D. CRITICAL RESOURCE SCHEDULING

Scheduling resources to get maximum performance, is crucial. Appendix B presents a resource scheduling problem and solution developed by CDR. M. Sarigul-Klijn of the

Alameda, CA., NADEP. In this problem, the conventional Critical Path Method (CPM) results in completion of the last aircraft 40 days later than CRS. CRS, developed from ToC, shows that, much of the schedule compliance problem occurs from the variation of the critical resource's TAT. The difference between the two methods is CPM first determines the longest path to complete one aircraft, CPM then labels it the critical path and schedules everything else around it. CRS, like CPM, first identifies the work breakdown structure, process time, and resources required. Unlike CPM, CRS identifies the critical resource, schedules all its requirements, then has all other resources scheduled subservient to it. Since process times are exponential, inductions should use worst case flow times from induction to the constraint. Using worst case flow times prevents the constraint from starving for work and builds the time buffer previously mentioned.

CPM with Project Evaluation and Review Techniques (PERT) looks not only at the critical path, but also makes a probabilistic estimate of project completion by taking in consideration early and late times of the servers' service time [Ref.7:p.8]. However, CPM/PERT does not look at

server's service rate restrictions. It assumes each server has an infinite service rate. This assumption in CPM/PERT causes AIP to increase, queue time to increase, and schedule compliance to decrease as process variation occurs at the unknown constraint.

Unfortunately, CPM/PERT is the only computer supported scheduling tool available to the Naval Aviation Depot (NADEP). To properly synchronize the remanufacturing line, CRS must be the scheduling tool. CRS is not presently available as a commercial software package.

E. DBR STRATEGY AS APPLIED TO AIRCRAFT SCHEDULING

1. Corporate Level

Initially, to reduce high AIP, inductions should be at a constant rate and less than the production rate of the remanufacturing line. The flow of inductions should be smooth so the NADEP has a deterministic rate to work with. Once the AIP is reduced, the induction rate should equal the production rate of the constraint shop.

2. Depot Level

The depot level manager should use the queuing charts to monitor the TAT and AIP of each division constraint and non-constraint. A full matrix organization will allow better management of the resources. It will allow for resources to quickly move to a developing constraint to increase the service rate and hold PU down to one or lower. If PU is above one then the server is not keeping up with the arrival rate. The policy for borrowing servers should be: first borrow from areas operating at the lowest PU and second borrow from the end of the production line where an induced constraint is not crucial. The depot level manager can give an accurate forecast of aircraft completion by using a queuing chart like Figure 15. Shops 31 through aircraft flight test are steady state and show a TAT of slightly more than one FY quarter. A completion date is then determined when the aircraft departs Shop 31. The effect of future aircraft inductions and completions on TAT and AIP are easily seen with the queuing charts. The planner simply extends the inducted and completion curves on Figure 15 with planned inductions and completions.

3. Divisional Level

The divisional manager should use CRS to develop the servers' time schedule and required PU level. From aircraft induction to delivery to the customer, CRS will accurately forecast remanufacturing flow.

VI CONCLUSION

Two characteristics of Naval Aviation Depot (NADEP) aircraft manufacturing operations, unknown requirements until disassembly and unscheduled workload, plague the depot managers in fulfilling the mission of providing a ready and controlled, cost effective remanufacturing resource. At the Naval Aviation Depot (NADEP), these two characteristics are actually constraints causing problems like high operating expense, high turnaround time (TAT), poor schedule compliance and less quality control.

The A-6 remanufacturing line at the NADEP, Alameda, CA., provided the research study area to investigate the application of queuing charts, Just-In-Time method, Theory of Constraints (ToC) principles, and Critical Resource Scheduling (CRS) with the goal of solving some of the problems of Department of Defense remanufacturing operations.

The A-6 line at Alameda consists of four processes: Standard Depot Level Maintenance (SDLM), SDLM aircraft composite wing rewing, SDLM aircraft wing modification 601, and EA-6B SDLM. Each of these processes are a series of

dependent events including acceptance, structural repair, modification, reassembly, paint, ground check, flight check, and delivery. This macro view of the A-6 line resembles the "I" logical manufacturing structure. The micro view of the line includes divisions that assign resources to rework component parts of the aircraft. The micro view resembles the "V/A" logical structure because the component parts are disassembled, reworked, then reassembled before delivery back to the aircraft.

Statistical fluctuations in the process time are very significant. These fluctuations are caused by the induction rate, variance in work content, late finds, engineering changes, material and personnel shortages, and high work in process.

Management at the NADEP, Alameda, believes in keeping labor producing to increase efficiency. This requires a large backlog of work, which results in high aircraft turnaround time, high operating costs, and losses for NADEP.

Queuing Theory defines the production philosophies of ToC. There are three basic elements to queuing systems: Customer, Server and Queue. There are three parts to the queuing system: Arrival Process, Service Process, and Departure

Process. There are several different measures of performance (MOP) in a queuing system which include: Arrival Rate, Service Rate, and Proportional Utilization (PU).

A Queuing Chart is a fluid approximation of the arrivals, queue departures and service departures of a single shop or a group of serial shops. Queuing Charts are a very robust and valuable tool to the depot manager as well as the divisional managers. The managers can now monitor important production MOP's, determine flow constraints, analyze the effect of lost service rates on flow, and make accurate forecasts of aircraft or component completion dates.

The study looked at the ToC approach, which has its root in Just-In-Time production methodology, and applied it to the planning and control of the NADEP remanufacturing operations. ToC accepts the existence of an unbalanced process and the most limited server is the constraint. ToC optimize the production process by synchronizing the remanufacturing flow with the Drum-Buffer-Rope method. The service rate of the constraint should set up the induction rate of aircraft to the whole system (The "Drum" beat). NAVAIR needs to know exactly what the service rate of the constraint is at all times (The "Rope"). The NADEP needs to strive for a Proportional

Utilization (PU) of one at the constraint by using time buffers to prevent it from starving for work (The "Buffer"). The NADEP should then elevate it by allocating additional resources to increase its service rate.

Scheduling the resources through Critical Resource Scheduling is far superior to the Critical Path Method with Project Evaluation and Review Techniques because it globally optimize the production schedule by incorporating the constraint of critical resources.

To fully implement the Critical Resource Scheduling method, several areas need follow-on research in the very near future:

- Develop a CRS scheduling computer program for use at the NADEP.
- Develop The "Rope", i.e., a communication network from the constraint shop to NAVAIR and material suppliers.
- Develop a three dimensional fluid approximation model to monitor dependent, serial <u>and</u> parallel servers together.

These methods require that the NADEPs change their planning and control methods. Change is never easy. Consider as an example the refusal of the United States to adopt metric measurements. The metric system of meters and kilograms has

intrinsic advantages over our English system of pounds and feet. Metric conversion factors are all factors of ten while English factors include 2, 4, 12, 16, 36, 144, and 5280 - among many others. So why does not the United States change to metric measurements? The answer is that we already know the old English measurements. Learning the easier metric measurements represents change and requires additional effort that we chose not to bother with.

In similar fashion, the NADEPs already know the conventional management methods of efficiencies, norms, variances, etc. The ToC methods are easier and promise substantial improvements, but require additional effort to learn.

The NADEPs should start using queuing charts and Critical Resource Scheduling measurements now. Eventually everyone will learn what they mean. As an example, most people have a feel for what "two liters" means thanks to the two liter soft drink bottle. Similarly with queuing charts, everyone will eventually understand the ToC measurements.

FINAL THOUGHTS

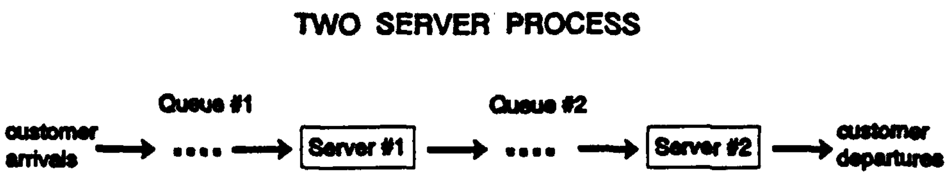
The methods proposed in this thesis will work only if relevant data is used. There was a large railroad company in the West which only measured trip times. It did not measure actual departure or arrival times. The company was happy if a trip took one hour, as it should, even though the train left five hours late and arrived five hours late. Its customers were not happy and were moving their business to trucking companies. The NADEPs do not measure actual arrival dates, instead they use induction dates. The difference in the most part are a few days or a few weeks, but examples of a few years exist. Similarly NADEPs do not measure actual departure dates. Instead they use sell or RFD (ready for delivery) dates. The difference in most cases is a few days or a few weeks but examples exist of several months.

Without relevant data, the methods proposed here simply will not work. IF one measures an organization in a dysfunctional way, then one can expect dysfunctional behavior and results.

APPENDIX A

Lotus 123 Spreadsheet Template

This appendix outlines the construction of cumulative queuing charts using the Lotus 123 spreadsheet program. Consider the following example of a remanufacturing process consisting of two servers in series:



Referring to the above system, the following is typically documented in remanufacturing:

- Arrival time into Queue #1 ($A(t)$).
- Arrival time into Queue #2 ($D_{S1}(t)$).
- Departure time from Server #2 ($D_{S2}(t)$).

The arrival times into Server #1 and Server #2 are typically not documented in remanufacturing. The parts might have arrived at the service shop but are not actually being serviced yet because of an in-shop backlog, they are instead sitting on shelves or in bins awaiting service to begin. With

this in mind, queue time is generally very large when compared to service time. Queue time is composed not only of time awaiting for parts, equipment, tools, and people, but also holidays and non-productive hours. Non-productive hours are hours when the NADEP is not working on customers that are inprocess. The NADEP has 168 hours (7 days X 24hr/day) per week to process customers, however, typically no more than 40 hours (5 work days X 8hrs/work day) are utilized. Thus an aircraft component spends more time waiting at each workcenter than being naturally worked on. A 40 hour process turnaround time takes 5 days with at least 80 hrs of sitting on a shelf or in a bin vice less than 2 days where the component is on a work bench being processed 100 percent of the time. The Lotus 123 spreadsheet template depicts this comparative turnaround time visually for the production manager. Note that the time unit called day, can be either the 24 hour calendar day or the 8 hour work day. Using the 24 hour calendar day depicts the global productivity level of the NADEP while using a day unit equal to the standard work day filters out the naturally non-productive hours allowing for the chart to depict more symptomatic productivity problems.

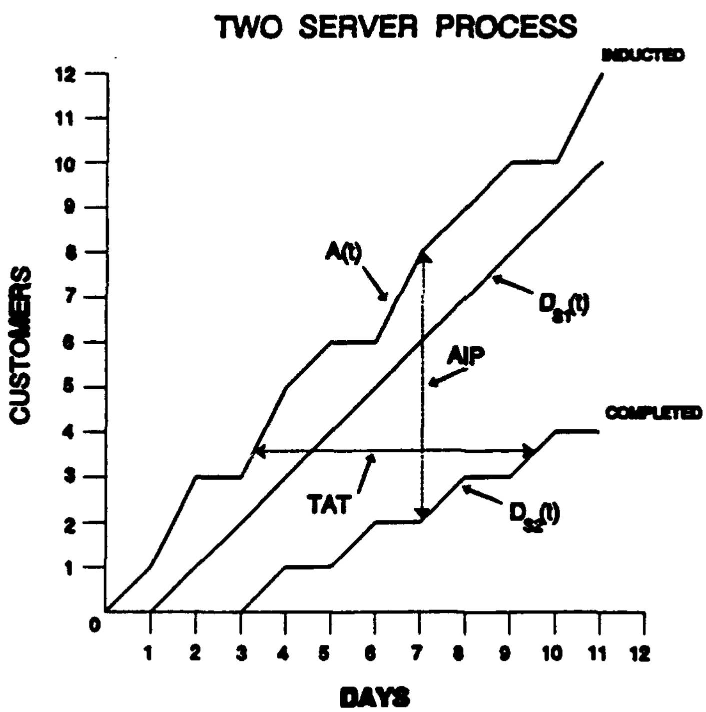
Suppose that customers in the two server example above arrive into the process from somewhere else in the remanufacturing system at a rate (A(t)) of zero to two customers per day. Server one has a service rate ($D_{s1}(t)$) of one customer per day and server two has a service rate ($D_{s2}(t)$) of one half customers per day.

TWO SERVER PROCESS

	A	B	C	D	E	F	G
1	DAY #	ARRIVALS QUEUE #1	CUMULATIVE ARRIVALS	ARRIVALS QUEUE #2	CUMULATIVE ARRIVALS	DEPARTING FROM SVC	CUMULATIVE DEPARTURE
2	0	0	0		0		0
3	1	1	1		0		0
4	2	2	3	1	1		0
5	3	0	3	1	2	0	0
6	4	2	5	1	4	1	1
7	5	1	6	1	5	0	1
8	6	0	6	1	6	1	2
9	7	2	8	1	7	0	2
10	8	1	9	1	8	1	3
11	9	1	10	1	9	0	3
12	10	0	10	1	10	1	4
13	11	2	12	1	11	0	4

Server arrival data is recorded in columns B and D, and process departure data is recorded in column F. The worksheet calculates cumulative arrival date in column C and E, and cumulative departure data in column G. The formula in cell C3 is: <u>+C2+B3</u>, in cell E3 is <u>+E2+D3</u>, and in cell G3 is <u>+G2+F3</u>. The formulas in the column C4 .. C13, E4 .. E13, and G4 .. G13 are copies of the formula in cell C3, E3 and G3 made by using

the /copy command. For example, after using the copy command the formula in cell C11 should be: <u>+ C10 + B11</u>. The following graph was formed by using the /graph command and by making the x-range equal to cells A2 .. A13, and the 3 y-ranges (A,B,C) equal to cells C2 .. C13, E2 .. E13, and G2 .. G13 respectfully. The following figure is the cumulative queuing chart for the two server system.



This queuing chart tells the manager: the arrival rate, A(t), of work is predictable over a two week period, something he can forecast with; server one has a steady output (probably an automated server) but is not keeping up with the arrival of

work into the process; server two has a predictable output when averaged over several weeks, however the output is less than server one so a large queue is developing. The manager has a problem that will not go away. The manager needs to get rid of the work-in-process (WIP) because it is costing him money and probably affecting the quality of his product. Doing away with the WIP will free up more liquid assets to invest in additional resources to increase the throughput of the servers. Once the WIP is gone, the manager can forecast the servers required throughput ensuring the proportional utilization is less than one to allow for surges in A(t).

APPENDIX B

Aircraft Resource Planning Problem

A typical aircraft resource planning problem is presented to demonstrate the advantage of Critical Resource Scheduling (CRS) over the more widely accepted Critical Path Method (CPM) of work scheduling. In the problem, three aircraft (A, B, C) use three resources (X shop, Y shop, Z shop) to complete an individual rework sequence. Each of the three aircraft rework sequences will use each of the three shops to complete the respective sequence schedule but in different amounts. One shop can work on only one aircraft at a time. The flow times vary because of the different work content required in each aircraft. The shops currently have no work in process and overtime, weekend working, etc., is already accounted for. Figure 1 specifies the shops, flow times, and sequence of operation. The customer wants to know the induction schedule, the ready for delivery (RFD) dates and the minimum time to complete his three aircraft.

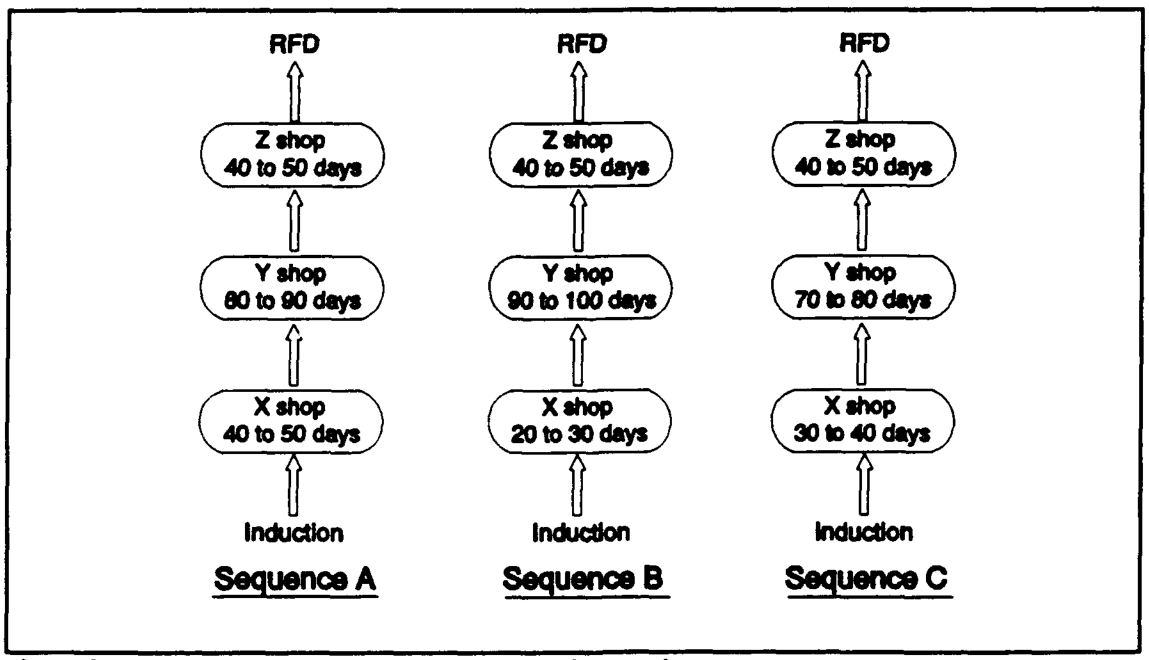


Figure 1: Aircraft resource planning sequence of operations.

Determine a schedule producing the minimum and maximum turn-around-time (TAT) for each aircraft and the minimum total time to complete all three rework sequences.

Solutions

CRITICAL PATH METHOD

1. The first step in CPM is identify the work breakdown. The work breakdown consists of the resources required, the order in which they are required, and the process time. This step has already been done and is shown as Figure 1 in the problem.

2. Determine the critical path. CPM first determines and labels the longest path in time as the critical path. For aircraft sequence A, X shop needs 40 to 50 days, Y shop 80 to 90 days, and Z shop 40 to 50 days for a total of 180 to 190 days. Similarly, sequence B needs a total of 150 to 180 days and sequence C needs a total of 140 to 170 days. Sequence A has the most days and is the critical path. Sequence B is the next most critical path and C is the least.

3. Build the schedule around the critical path. The schedule for the critical path is first written. The other parallel paths have slack time relative to the critical path resulting in the option of an early or late starting date, relative to the critical path. An important point and not always clearly understood is that CPM assumes infinite resources.

4. In actual practice, CPM results in a schedule which cannot be executed because it assumes infinite resources. Figure 2 presents a schedule which obeys the priority scheme of CPM and considers the available resources. It shows that the competition for the resources X, Y and Z are responsible for pushing the earliest completion of the three aircraft to 320 days and the latest completion to 370 days. On the right side of the figure is turnaround time (TAT), that is, the completion date minus the induction date. Aircraft sequence B and C experience significant queue time.

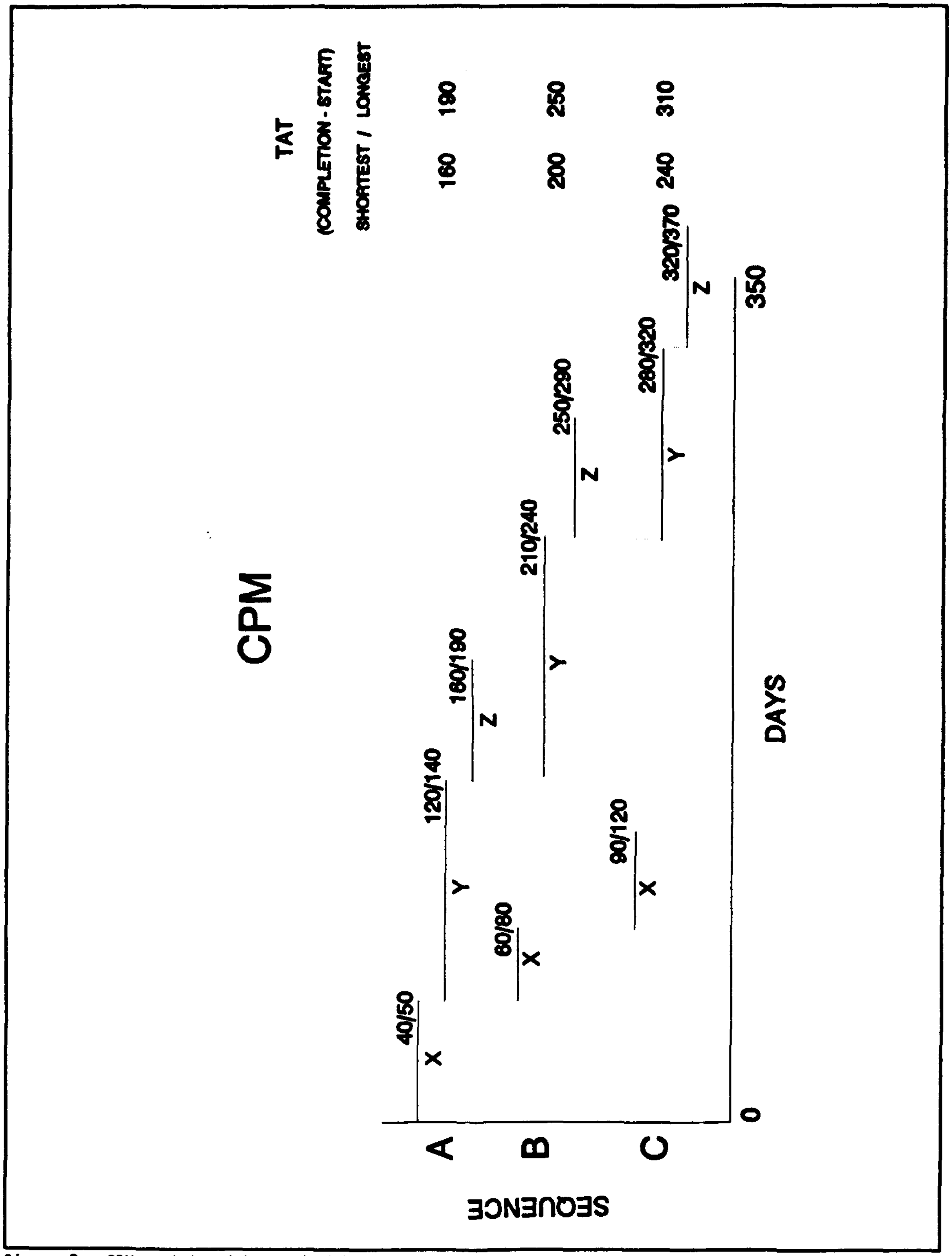


Figure 2: CPM work breakdown schedule.

CRITICAL RESOURCE SCHEDULING

1. The first step in CRS is to identify the work breakdown. This first step is the same as in CPM.

2. Next determine the critical resource. Shop X needs 40 to 50 days to complete aircraft sequence A, 20 to 30 days for B, 30 to 40 days for C, for a total of 90 to 120 days to complete all three aircraft. Similarly, the days for the Y and Z shop can be totalled up. The Y shop needs a total of 240 to 270 days to complete all three aircraft. Finally, the Z shop needs 120 to 150 days for all three. The Y shop is the critical resource because it needs the most days to complete the aircraft. The Z shop is the next most critical resource and the X shop is the least critical resource.

3. Build the schedule around the critical resource. The idea here is to get the critical resource started as early as possible and then to schedule aircraft so that the critical resource is never idle. Shop X can complete aircraft B in the quickest time (20 to 30 days), so schedule it first. Schedule aircraft sequence C next because it takes less days (70 to 80) through the Y shop than sequence A (80 to 90 days). This choice keeps the next most critical resource, Shop Z, busy the earliest. Schedule aircraft A last.

4. Backward schedule from critical resource (Y shop) start date to determine the induction dates for the aircraft. Use the most optimism schedule (least number of days) for the Y shop and the most pessimist schedule (most number of days) for the X shop when determining the induction schedule. This will ensure that the Y shop is never starved for work. Similarly, forward schedule from the Y shop completion days for Ready to Delivery dates. Use the most pessimist days throughout all shops to guarantee delivery on a specific date. Figure 3 presents the CRS schedule. All aircraft experience minimum queue time with the CRS schedule. Turnaround time is significantly better than the CPM schedule.

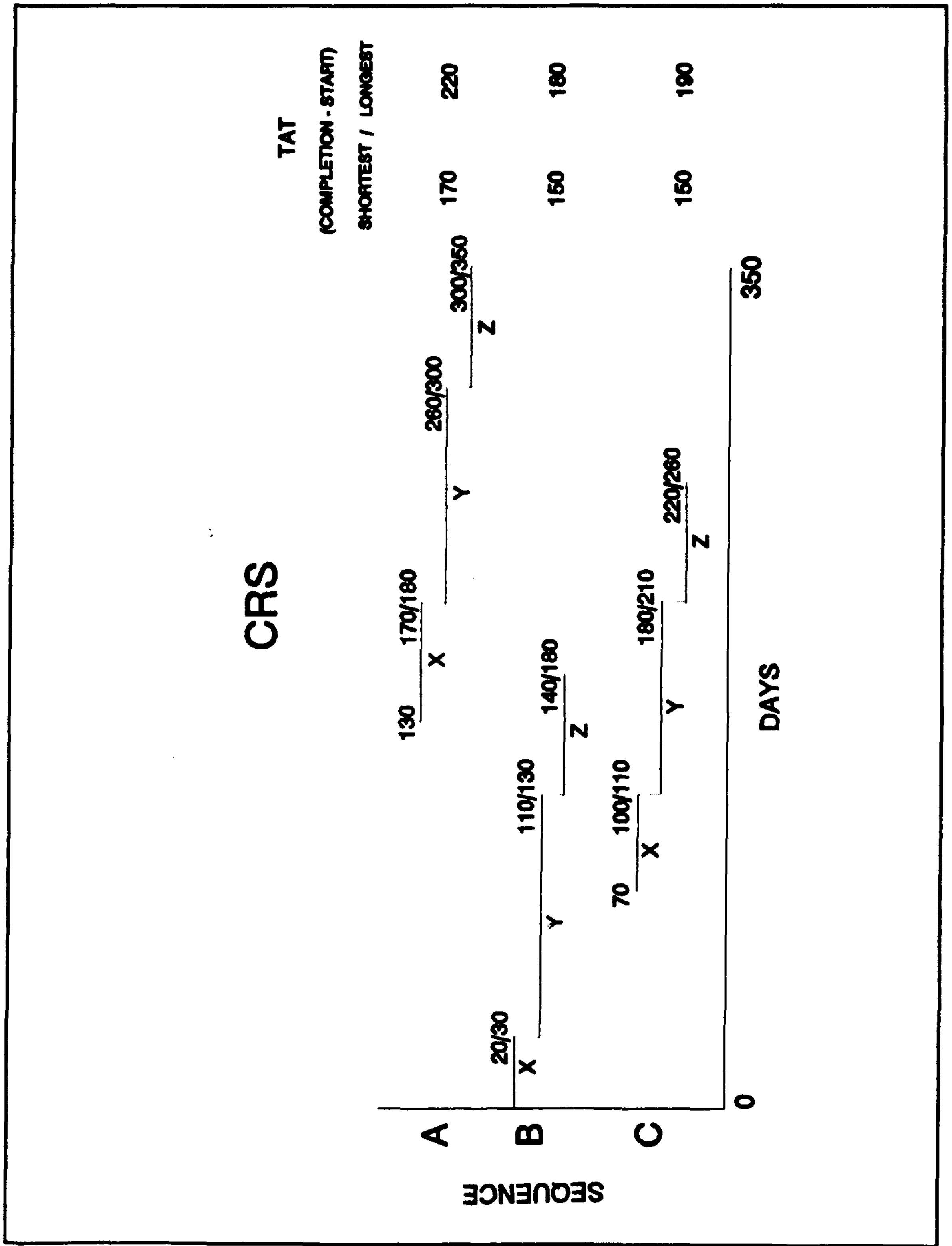


Figure 2: CRS work breakdown schedule.

What is the guaranteed delivery time for the three aircraft?

CRS = 350 days; CPM = 370 days

In conclusion Critical Resource Scheduling:

- Works well with a reservation system to get the customer inducted at the specified induction date.
- Does not cause the customer to have to experience any significant queue time.
- Does not cause the remanufacturer to have to endure high work in process.
- Has built in time buffers allowing the system to flex with induction rates or shop productivity.

LIST OF REFERENCES

1. Appel, Jeffrey S., "The Naval Aviation Depot," Research Report, 1989.

2. Hall, Randolph W., Queueing Methods for Services and Manufacturing, Prentice-Hall, Inc., New Jersey, 1991.

3. Blackstone, John H., Capacity Management, South-Western Publishing, 1989.

4. Fogarty, Donald W., Blackstone, John H., and Hoffman, Thomas R., Production and Inventory Management, South-Western Publishing, 1991.

5. Goldratt, Eliyahu M., and Cox, Jeff, The Goal: A Process of Ongoing Improvement, North River Press, Inc., 1986.

6. Cox, James F., and Finch, Byron, Effective Placement of Manufacturing Inventories, American Production and Inventory Control Society, Conference Proceedings, 1989.

7. NAVAIR, Theory of Constraints Development Strategy, MRP II Implementation Planning and Support Contract Number: 09076-008-120, Implementation Report 1990.

INITIAL DISTRIBUTION LIST

1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22304-6145	2
2.	Library, Code 52 Naval Postgraduate School Monterey, California 93943-5002	2
3.	Prof. Shu Liao (Code AS/Lc) Naval Postgraduate School Monterey, California 93943-5000	1
4.	Commanding Officer COMNAVAIRWARCENWPNDIV ATTN: CDR Marti Sarigul-Klijn Pt. Mugu, California 93042	4
5.	LCDR James M. Beck 12751 Calle De Las Rosas San Diego, California 92129	2